

Convex analysis for composite functions without K-convexity

Juan Pablo Vielma

Google Research, jvielma@google.com

Abstract. Composite functions have been studied for over 40 years and appear in a wide range of optimization problems. Convex analysis of these functions focuses on (i) conditions for convexity of the function based on properties of its components, (ii) formulas for the convex conjugate of the function based on those of its components and (iii) chain-rule-like formulas for the sub-differential of the function. To the best of our knowledge all existing results on this matter are based on the notion of K-convexity of functions where K is a closed convex cone. These notions can be considered elementary when K is the non-negative orthant, but otherwise may limit the accessibility of the associated results. In this work we show how standard results on perturbation function based duality can be used to recover and extend existing results without the need for K-convexity. We also provide a detailed comparison with K-convexity based results and show that, while K-convexity is not needed for the convex analysis of composite functions, it certainly aids in verifying the conditions associated with our proposed results. Finally, we provide necessary conditions for the conjugate and sub-differential formulas that are less restrictive than those required by the existing results.

Key words: composite function, convex conjugate, subdifferential, K-convexity

1. Introduction For $g : \mathbb{R}^m \rightarrow \mathbb{R}$ and $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ consider the composite function $g \circ F : \mathbb{R}^n \rightarrow \mathbb{R}$ given by

$$g \circ F(x) = g(F(x)) \quad \forall x \in \mathbb{R}^n.$$

As noted in [6], functions of the form $g \circ F$ have been studied for over 40 years and appear as objective functions for a wide range of optimization problems. In particular, the case where $g \circ F$ is convex has been extensively studied (e.g. [3–8, 11]).

Some natural questions in the convex case include (i) conditions for convexity of $g \circ F$ based on properties of g and F , (ii) formulas for the convex conjugate of $g \circ F$ based on the conjugates of g and F , and (iii) chain-rule-like formulas for the sub-differential of $g \circ F$. A standard sufficient condition for convexity of $g \circ F$ when F is a real-valued function (i.e. $m = 1$) is monotonicity of g and convexity of F . A first extension of these conditions to vector-valued F (i.e. $m > 1$) is to consider component-wise convexity of F and component-wise monotonicity of g . As shown in [9], in addition to convexity of $g \circ F$, these component-wise conditions also yield a relatively simple formula for the convex conjugate of $g \circ F$.

The notions of component-wise convexity of F and component-wise monotonicity are based on the partial order in \mathbb{R}^m induced by the closed convex cone \mathbb{R}_+^m (i.e. for $x, y \in \mathbb{R}^m$ we have $x \leq y$ if and

only if $y - x \in \mathbb{R}_+^m$). Hence, it is natural to extend these concepts by considering an arbitrary closed convex cone K to introduce the notions of K -convexity and K -monotonicity (e.g. see Section 4). To the best of our knowledge, all existing studies of $g \circ F$ for the convex case with $m > 1$ (including [3–8, 11]) are based on these notions of K -convexity and K -monotonicity. In particular, combining these notions with perturbation-function based dualization (e.g. [15, Section 11.H]) can result in fairly general answers to all three questions above even in the infinite-dimensional setting (e.g. [2, 5]).

It can be argued that the need for beyond-elemental convex analysis concepts such as perturbation-functions and K -convexity can limit the accessibility of the results. Motivated by this potential accessibility issue, [6] focused on the finite dimensional case without directly applying the perturbation-function machinery to improve the accessibility of the associated result development and proofs. However, the statements of the associated results in [6] (e.g. necessary conditions and formulas) still require the notions of K -convexity and K -monotonicity. For this reason, we take a different approach and focus on avoiding the notions of K -convexity and K -monotonicity altogether to obtain results whose statements only require elemental convex analysis concepts. That is, our focus is primarily to improve the accessibility for readers that seek to apply the results without necessarily looking at the proofs. Similarly to [2, 5], the proofs of our results are based on perturbation-function based dualization, but in contrast to [2, 5] we only use a version that does not require the notions of K -convexity or K -monotonicity. Hence, we also get some accessibility improvements on the proofs themselves.

One additional advantage of avoiding K -convexity is that we can focus on additional properties of the functions g and F (e.g. polyhedrality) which have been exploited by standard perturbation-function based dualization. This allows us to obtain necessary conditions for the conjugate and sub-differential formulas that are less restrictive than those required by the existing results in [3–8, 11]. However, a potential disadvantage is that some of these necessary conditions may be harder to verify than those based on K -convexity. For instance, one of these conditions requires convexity of a perturbation function that, while simple to state, could be hard to verify. Fortunately, this joint convexity is implied by some standard K -convexity conditions. For this reason we also provide a detailed comparison between our results and those based on K -convexity. In particular, we show how the results based on K -convexity can be more restrictive, but can aid in verifying the less restrictive conditions we consider.

The remainder of this paper is structured as follows. In Section 1.1 we introduce some notation and assumptions. Then, in Section 2 we review the perturbation-function based dualization results we will use for our main results, which we present in Section 3. After that, we compare our main results with those based on K -convexity in Section 4. Finally, some omitted proofs are presented in Section 5 and Appendix A.

1.1. Notation and assumptions To consider extended real valued functions we let $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$. For simplicity we use \mathbb{R}^p for some $p \geq 1$ with inner product denoted by $\langle \cdot, \cdot \rangle$ as our vector spaces, but all our results also hold for Euclidean spaces through standard vectorization transformations (e.g. the space of $n \times n$ symmetric matrices can be *unrolled* to \mathbb{R}^p with $p = n(n+1)/2$). For a set $S \subseteq \mathbb{R}^p$ we let $\text{rint}(S)$ be its relative interior and $\text{cl}(S)$ be its closure. For a function $h : \mathbb{R}^p \rightarrow \overline{\mathbb{R}}$ we let $\text{dom}(h) := \{x \in \mathbb{R}^p : h(x) < +\infty\}$ be the effective domain of h , $\text{epi}(h) := \{(x, \alpha) \in \mathbb{R}^p \times \mathbb{R} : h(x) \leq \alpha\}$ be the epigraph of h , $\text{cl}(h) : \mathbb{R}^p \rightarrow \overline{\mathbb{R}}$ be the function such that $\text{epi}(\text{cl}(h)) = \text{cl}(\text{epi}(h))$, and $h^* : \mathbb{R}^p \rightarrow \overline{\mathbb{R}}$ be the convex conjugate of h given by

$$h^*(v) := \sup_{x \in \mathbb{R}^p} \{\langle v, x \rangle - h(x)\} \quad \forall v \in \mathbb{R}^p.$$

We say a function $h : \mathbb{R}^p \rightarrow \overline{\mathbb{R}}$ is *convex* if $\text{epi}(h)$ is convex, is *closed* if $\text{epi}(h)$ is closed, and is *proper* if $h \not\equiv +\infty$ and $h(x) > -\infty$ for all $x \in \mathbb{R}^p$. We also let

$$\begin{aligned} \Gamma(\mathbb{R}^p) &:= \left\{ h : \mathbb{R}^p \rightarrow \overline{\mathbb{R}} : h \text{ is proper and convex} \right\} \\ \Gamma_0(\mathbb{R}^p) &:= \{ h \in \Gamma(\mathbb{R}^p) : h \text{ is closed} \}. \end{aligned}$$

For a closed convex cone $K \subseteq \mathbb{R}^p$ we let $K^\circ := \{v \in \mathbb{R}^p : \langle x, v \rangle \leq 0 \quad \forall x \in K\}$.

Finally, we use the following definitions to simplify some of our statements (e.g. to avoid pathological cases of $\arg \min$ without having to restrict its definition to proper functions).

DEFINITION 1.1. For $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ we let

$$\begin{aligned} \arg \min_{x \in \mathbb{R}^n} h(x) &:= \left\{ \bar{x} \in \mathbb{R}^n : h(\bar{x}) \in \mathbb{R}, \quad h(\bar{x}) = \inf_{x \in \mathbb{R}^n} h(x) \right\} \\ \arg \max_{x \in \mathbb{R}^n} h(x) &:= \left\{ \bar{x} \in \mathbb{R}^n : h(\bar{x}) \in \mathbb{R}, \quad h(\bar{x}) = \sup_{x \in \mathbb{R}^n} h(x) \right\}. \end{aligned}$$

If h is convex and $\bar{x} \in \mathbb{R}^n$ we let

$$\partial h(\bar{x}) := \begin{cases} \{v \in \mathbb{R}^n : h(x) \geq h(\bar{x}) + \langle v, (x - \bar{x}) \rangle \quad \forall x \in \mathbb{R}^n\} & h(\bar{x}) \in \mathbb{R} \\ \emptyset & \text{o.w.} \end{cases}.$$

2. Perturbation based Lagrangian duality For $f \in \Gamma_0(\mathbb{R}^n \times \mathbb{R}^m)$ [15, Section 11.H] presents properties of the primal-dual pair

$$\text{opt}_P := \inf_{x \in \mathbb{R}^n} f(x, 0) \quad (2.1a)$$

$$\text{opt}_D := \sup_{y \in \mathbb{R}^m} -f^*(0, y). \quad (2.1b)$$

For instance, [15, Theorem 11.39] shows that $\text{opt}_P \geq \text{opt}_D$ and gives conditions for this to hold at equality.

However, the results in [15, Section 11.H] also imply properties of the slightly more general primal-dual pairs

$$\text{opt}_P(\bar{u}, \bar{v}) := \inf_{x \in \mathbb{R}^n} f(x, \bar{u}) - \langle \bar{v}, x \rangle \quad (2.2a)$$

$$\text{opt}_D(\bar{u}, \bar{v}) := \sup_{y \in \mathbb{R}^m} \langle \bar{u}, y \rangle - f^*(\bar{v}, y) \quad (2.2b)$$

whose parametrization on $(\bar{v}, \bar{u}) \in \mathbb{R}^n \times \mathbb{R}^m$ can be used to study the conjugate functions (e.g. $\text{opt}_P(0, \bar{v}) = -(f(\cdot, 0))^*(\bar{v})$). We now summarize these properties and connect them to the generalized nonlinear programming framework of [13]. While all these results follow from [15, Section 11.H] we give detailed proofs in Appendix A for completeness. We begin with the basic definitions of Lagrangian function, primal and dual value functions, and primal and dual optimal sets.

DEFINITION 2.1. For $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \bar{\mathbb{R}}$, let $l : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \bar{\mathbb{R}}$ be the *Lagrangian function* given by

$$l(x, y) := -(f(x, \cdot))^*(y). \quad (2.3)$$

In addition, for $\bar{v} \in \mathbb{R}^n$, and $\bar{u} \in \mathbb{R}^m$, let $p_{\bar{v}} : \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$ be the *primal value function* and $q_{\bar{u}} : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be the *negative of the dual value function* given by

$$p_{\bar{v}}(\bar{u}) := \inf_{x \in \mathbb{R}^n} \{f(x, \bar{u}) - \langle \bar{v}, x \rangle\} = -\sup_{x \in \mathbb{R}^n} \{\langle \bar{v}, x \rangle - f(x, \bar{u})\} \quad (2.4a)$$

$$q_{\bar{u}}(\bar{v}) := \inf_{y \in \mathbb{R}^m} \{f^*(\bar{v}, y) - \langle \bar{u}, y \rangle\} = -\sup_{y \in \mathbb{R}^m} \{\langle \bar{u}, y \rangle - f^*(\bar{v}, y)\}. \quad (2.4b)$$

Finally, let $P(\bar{u}, \bar{v}) \subseteq \mathbb{R}^n$ be the *primal optimal set* and $Q(\bar{u}, \bar{v}) \subseteq \mathbb{R}^m$ be the *dual optimal set* given by

$$\begin{aligned} P(\bar{u}, \bar{v}) &:= \arg \min_{x \in \mathbb{R}^n} \{f(x, \bar{u}) - \langle \bar{v}, x \rangle\} = \arg \max_{x \in \mathbb{R}^n} \{\langle \bar{v}, x \rangle - f(x, \bar{u})\} \\ Q(\bar{u}, \bar{v}) &:= \arg \min_{y \in \mathbb{R}^m} \{f^*(\bar{v}, y) - \langle \bar{u}, y \rangle\} = \arg \max_{y \in \mathbb{R}^m} \{\langle \bar{u}, y \rangle - f^*(\bar{v}, y)\}. \end{aligned}$$

2.1. General weak duality and optimality conditions The following theorem gives various flavors of weak duality and some optimality conditions that do not require any requirements on f .

THEOREM 2.1. *For any $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$ and $(\bar{v}, \bar{u}) \in \mathbb{R}^n \times \mathbb{R}^m$ we have that $q_{\bar{u}}$ is convex,*

$$f^*(\bar{v}, \bar{y}) - \langle \bar{u}, \bar{y} \rangle = \sup_{x \in \mathbb{R}^n} \{\langle \bar{v}, x \rangle - \langle \bar{u}, \bar{y} \rangle - l(x, \bar{y})\} \quad \forall \bar{y} \in \mathbb{R}^m$$

and $q_{\bar{u}}(\bar{v}) := \inf_{y \in \mathbb{R}^m} \sup_{x \in \mathbb{R}^n} \{\langle \bar{v}, x \rangle - \langle \bar{u}, y \rangle - l(x, y)\}$.

In addition,

$$\text{opt}_P(\bar{u}, \bar{v}) = p_{\bar{v}}(\bar{u}) \geq \text{cl co } p_{\bar{v}}(\bar{u}) \geq -q_{\bar{u}}(\bar{v}) = \text{opt}_D(\bar{u}, \bar{v}), \quad (2.5)$$

$$-\text{opt}_P(\bar{u}, \bar{v}) = -p_{\bar{v}}(\bar{u}) \leq \text{cl } q_{\bar{u}}(\bar{v}) \leq q_{\bar{u}}(\bar{v}) = -\text{opt}_D(\bar{u}, \bar{v}). \quad (2.6)$$

Finally, for any $(\bar{v}, \bar{u}) \in \mathbb{R}^n \times \mathbb{R}^m$ and $(\bar{x}, \bar{y}) \in \mathbb{R}^n \times \mathbb{R}^m$

$$\bar{x} \in P(\bar{u}, \bar{v}), \quad \bar{y} \in Q(\bar{u}, \bar{v}), \quad \text{and} \quad p_{\bar{v}}(\bar{u}) = -q_{\bar{u}}(\bar{v}) \quad (2.7)$$

is equivalent to

$$f(\bar{x}, \bar{u}) - \langle \bar{v}, \bar{x} \rangle = l(\bar{x}, \bar{y}) + \langle \bar{u}, \bar{y} \rangle - \langle \bar{v}, \bar{x} \rangle = \langle \bar{u}, \bar{y} \rangle - f^*(\bar{v}, \bar{y}) \in \mathbb{R}. \quad (2.8)$$

→ See page 30 for the proof.

In particular, we have that both (2.5) and (2.6) yield weak duality $\text{opt}_P \geq \text{opt}_D$ for primal-dual pair (2.1), but (2.5) allows for considering parametric limits on the primal value function p_v while (2.6) allows the same on the (negative of) the dual value function q_u . Through the equivalence between (2.7) and (2.8) Theorem 2.1 yields optimality conditions based on the Lagrangian function l . However, to obtain true *saddle-point* conditions we need some convexity conditions on f as described in the following proposition.

PROPOSITION 2.1. *For any proper $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$ such that $f(x, u)$ is closed and convex in u , and for any $(\bar{v}, \bar{u}) \in \mathbb{R}^n \times \mathbb{R}^m$ we have*

$$f(\bar{x}, \bar{u}) - \langle \bar{v}, \bar{x} \rangle = \sup_{y \in \mathbb{R}^m} \{l(\bar{x}, y) + \langle \bar{u}, y \rangle - \langle \bar{v}, \bar{x} \rangle\} \quad \forall \bar{x} \in \mathbb{R}^n$$

and $p_{\bar{v}}(\bar{u}) := \inf_{x \in \mathbb{R}^n} \sup_{y \in \mathbb{R}^m} \{l(x, y) + \langle \bar{u}, y \rangle - \langle \bar{v}, x \rangle\}$.

In addition, (2.8) is equivalent to

$$\bar{x} \in \arg \min_{x \in \mathbb{R}^n} \{l(x, \bar{y}) + \langle \bar{u}, \bar{y} \rangle - \langle \bar{v}, x \rangle\}, \quad (2.9a)$$

$$\bar{y} \in \arg \max_{y \in \mathbb{R}^m} \{l(\bar{x}, y) + \langle \bar{u}, y \rangle - \langle \bar{v}, \bar{x} \rangle\}. \quad (2.9b)$$

→ See page 31 for the proof.

2.2. Convex strong duality The following theorem gives various flavors of strong duality and characterizations of the optimal sets under convexity of f .

THEOREM 2.2. *If $f \in \Gamma(\mathbb{R}^n \times \mathbb{R}^m)$ and $f(x, u)$ is closed in u , then $l(x, y)$ is convex in x and concave in y and (2.9) is equivalent to*

$$0 \in \partial(l(\cdot, \bar{y}) + \langle \bar{u}, \bar{y} \rangle - \langle \bar{v}, \cdot \rangle)(\bar{x}) \quad (2.10a)$$

$$0 \in \partial(-l(\bar{x}, \cdot) - \langle \bar{u}, \cdot \rangle + \langle \bar{v}, \bar{x} \rangle)(\bar{y}). \quad (2.10b)$$

In addition, if $f \in \Gamma_0(\mathbb{R}^n \times \mathbb{R}^m)$ we have

1. **Optimal sets:** For all $(\bar{v}, \bar{u}) \in \mathbb{R}^n \times \mathbb{R}^m$ we have that $p_{\bar{v}}$ and $q_{\bar{u}}$ are convex, and

$$P(\bar{u}, \bar{v}) = \partial q_{\bar{u}}(\bar{v}) \quad (2.11a)$$

$$Q(\bar{u}, \bar{v}) = \partial p_{\bar{v}}(\bar{u}). \quad (2.11b)$$

2. **Primal-asymptotic strong duality:** If $p_{\bar{v}}$ is proper, then $p_{\bar{v}}(\bar{u}) \geq \text{cl } p_{\bar{v}}(\bar{u}) = -q_{\bar{u}}(\bar{v})$ for all $\bar{u} \in \mathbb{R}^m$.

3. **Primal-uniform strong duality:** If $p_{\bar{v}}$ is proper and closed, then $p_{\bar{v}}(\bar{u}) = -q_{\bar{u}}(\bar{v})$ for all $\bar{u} \in \mathbb{R}^m$.

4. **Dual-asymptotic strong duality:** If $q_{\bar{u}}$ is proper, then $-p_{\bar{v}}(\bar{u}) = \text{cl } q_{\bar{u}}(\bar{v})$ for all $\bar{v} \in \mathbb{R}^n$.

5. **Dual-uniform strong duality:** If $q_{\bar{u}}$ is proper and closed, then $-p_{\bar{v}}(\bar{u}) = q_{\bar{u}}(\bar{v})$ for all $\bar{v} \in \mathbb{R}^n$.

→ See page 32 for the proof.

Theorem 2.2 implicitly contains three classes of qualification conditions beyond the baseline $f \in \Gamma_0(\mathbb{R}^n \times \mathbb{R}^m)$.

First, requiring $p_{\bar{v}}$ or $q_{\bar{u}}$ to be proper yield asymptotic versions of strong duality where a zero duality gap is achieved through taking limits on the primal or dual perturbations \bar{u} or \bar{v} .

Second, requiring $p_{\bar{v}}$ or $q_{\bar{u}}$ to be proper and closed yields non-asymptotic versions of strong duality where zero duality gap is achieved for all perturbations.

Third, non-emptiness of the sub-differentials in (2.11) yields attainment of the primal or dual optimal values.

Describing these qualification conditions through properties of $p_{\bar{v}}$ and $q_{\bar{u}}$ instead of more typical conditions (e.g. Slater-like conditions) allows for generic and possibly ad-hoc conditions that exploit the structure of particular instantiations of these functions (and indirectly of f). However, these generic conditions are also implied by more typical conditions, which can be a better trade-off between applicability and simplicity. Corollary 2.1 and Corollary 2.2 below gives a Slater-like conditions that we will use in our study of composite functions.

DEFINITION 2.2. $h : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ is *piecewise linear-quadratic* (PWLQ) if and only if

$$h(x) = \begin{cases} \frac{1}{2} \langle x, A^i x \rangle + \langle q^i, x \rangle - u_i & x \in P_i, i \in I \\ +\infty & \text{o.w.} \end{cases},$$

where I is finite and for each $i \in I$, P_i is a polyhedron, $A^i \in \mathbb{R}^{n \times n}$, $q^i \in \mathbb{R}^n$ and $u_i \in \mathbb{R}$.

COROLLARY 2.1. [Primal Slater Condition] Let $f \in \Gamma_0(\mathbb{R}^n \times \mathbb{R}^m)$, $\bar{v} \in \mathbb{R}^n$, $\bar{u} \in \mathbb{R}^m$, and $p_{\bar{v}} : \mathbb{R}^m \rightarrow \bar{\mathbb{R}}$, $q_{\bar{u}} : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$ be the functions given by (2.4). Then,

- $q_{\bar{u}}$ is proper if $\bar{u} \in \text{dom}(p_0)$,
- $q_{\bar{u}}$ is closed if $\bar{u} \in \text{rint dom}(p_0)$, or $\bar{u} \in \text{dom}(p_0)$ and f is PWLQ, and
- $\partial p_{\bar{v}}(\bar{u}) \neq \emptyset$ if $\bar{u} \in \text{rint dom}(p_0)$ and $p_{\bar{v}}(\bar{u}) > -\infty$, or $\bar{u} \in \text{dom}(p_0)$ and f is PWLQ,

In particular, if either $\bar{u} \in \text{rint dom}(p_0)$, or $\bar{u} \in \text{dom}(p_0)$ and f is PWLQ, then for any $\bar{v} \in \mathbb{R}^n$ we have

$$\exists \bar{y} \in \mathbb{R}^m \quad s.t \quad \langle \bar{u}, \bar{y} \rangle - f^*(\bar{v}, \bar{y}) = -q_{\bar{u}}(\bar{v}) = p_{\bar{v}}(\bar{u}) \in \mathbb{R} \cup \{-\infty\}. \quad (2.12)$$

→ See page 35 for the proof.

Note that (2.12) implies that strong duality holds for primal-dual pair (2.2) and the dual optimal value is attained if finite.

COROLLARY 2.2. [Dual Slater Condition] Let $f \in \Gamma_0(\mathbb{R}^n \times \mathbb{R}^m)$, $\bar{v} \in \mathbb{R}^n$, $\bar{u} \in \mathbb{R}^m$, and $p_{\bar{v}} : \mathbb{R}^m \rightarrow \bar{\mathbb{R}}$, $q_{\bar{u}} : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$ be the functions given by (2.4). Then,

- $p_{\bar{v}}$ is proper if $\bar{v} \in \text{dom}(q_0)$,
- $p_{\bar{v}}$ is closed if $\bar{v} \in \text{rint dom}(q_0)$, or $\bar{v} \in \text{dom}(q_0)$ and f is PWLQ, and
- $\partial q_{\bar{u}}(\bar{v}) \neq \emptyset$ if $\bar{v} \in \text{rint dom}(q_0)$ and $q_{\bar{u}}(\bar{v}) > -\infty$, or $\bar{v} \in \text{dom}(q_0)$ and f is PWLQ.

In particular, if either $\bar{v} \in \text{rint dom}(q_0)$, or $\bar{v} \in \text{dom}(q_0)$ and f is PWLQ, then for any $\bar{u} \in \mathbb{R}^m$ we have

$$\exists \bar{x} \in \mathbb{R}^n \quad s.t \quad f(\bar{x}, \bar{u}) - \langle \bar{v}, \bar{x} \rangle = p_{\bar{v}}(\bar{u}) = -q_{\bar{u}}(\bar{v}) \in \mathbb{R} \cup \{+\infty\}. \quad (2.13)$$

→ See page 35 for the proof.

Note that (2.13) implies that strong duality holds for primal-dual pair (2.2) and the primal optimal value is attained if finite.

2.3. Generalized nonlinear programming The generalized nonlinear programming (GNLP) framework of [13] follows from the results in [15, Section 11.H] by taking the specific function f described in the following lemma. A very simple extension of the GNLP framework allows us to cover composite functions in their full generality (see Lemma 3.1). Finally, we note that the GNLP framework has theoretical and practical implications that go beyond the scope of this paper (e.g. see [13, 14]).

LEMMA 2.1 ([13]). For $f_0 : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$, $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $g : \mathbb{R}^m \rightarrow \bar{\mathbb{R}}$ let $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \bar{\mathbb{R}}$ be given by

$$f(x, u) := f_0(x) + g(F(x) + u)$$

and $l : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$ be given by

$$l(x, y) := -(f(x, \cdot))^*(y).$$

Then $l(x, y) = f_0(x) + \langle y, F(x) \rangle - g^*(y)$ for all $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$.

3. Composite functions To cover composite functions in the full generality of [6] we only need to extend the GNLP framework of Section 2.3 to allow F to take values in $\mathbb{R}^m \cup \{+\infty_\bullet\}$ where $+\infty_\bullet$ is an *infinite element* for \mathbb{R}^m as follows.

DEFINITION 3.1. Let $+\infty_\bullet$ be an *infinite element* for \mathbb{R}^m . For $F : \mathbb{R}^n \rightarrow \mathbb{R}^m \cup \{+\infty_\bullet\}$ we let $\text{dom}(F) := \{x \in \mathbb{R}^n : F(x) \neq +\infty_\bullet\}$ and $\text{rge}(F) = F(\text{dom}(F))$.

Here, infinite element $+\infty_\bullet$ is intended to represent all values for F not in \mathbb{R}^m . For instance, we can redefine any $H : \mathbb{R}^n \rightarrow (\mathbb{R} \cup \{-\infty, +\infty\})^m$ to a function $F : \mathbb{R}^n \rightarrow \mathbb{R}^m \cup \{+\infty_\bullet\}$ by letting

$$F(x) := \begin{cases} H(x) & H(x) \in \mathbb{R}^m \\ +\infty_\bullet & \text{o.w.} \end{cases}.$$

In Definition 4.1, $+\infty_\bullet$ will correspond to a *largest element* for a specific ordering.

Lemma 2.1 then extends to the following lemma.

LEMMA 3.1. Let $f_0 : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$, $F : \mathbb{R}^n \rightarrow \mathbb{R}^m \cup \{+\infty_\bullet\}$ and $g : \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$ let $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$, $\langle y, F \rangle : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ and $l : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$ be given by

$$f(x, u) := \begin{cases} f_0(x) + g(F(x) + u) & x \in \text{dom}(F) \\ +\infty & \text{o.w.} \end{cases}, \quad (3.1a)$$

$$\langle y, F \rangle(x) := \begin{cases} \langle y, F(x) \rangle & x \in \text{dom}(F) \\ +\infty & \text{o.w.} \end{cases}, \quad (3.1b)$$

$$l(x, y) := -(f(x, \cdot))^*(y). \quad (3.1c)$$

Then

$$l(x, y) = f_0(x) + \langle y, F(x) \rangle - g^*(y), \quad (3.2a)$$

$$f^*(v, y) = (f_0 + \langle y, F \rangle)^*(v) + g^*(y). \quad (3.2b)$$

→ See page 24 for the proof.

Finally, applying the results from Section 2 to the function f from Lemma 3.1 we get the following characterization of the convex conjugate of composite functions.

THEOREM 3.1. *For $f_0 : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$, $F : \mathbb{R}^n \rightarrow \mathbb{R}^m \cup \{+\infty\}$ and $g : \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$, let $\langle y, F \rangle : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ given by (3.1b), $\rho : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be given by*

$$\rho(\bar{v}) := \inf_{y \in \mathbb{R}^m} (f_0 + \langle y, F \rangle)^*(\bar{v}) + g^*(y), \quad (3.3)$$

and $f_0 + g \circ F : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be given by

$$(f_0 + g \circ F)(x) := \begin{cases} f_0(x) + g(F(x)) & x \in \text{dom}(F) \\ +\infty & \text{o.w.} \end{cases}. \quad (3.4)$$

Then, ρ is convex, for any $\bar{v} \in \mathbb{R}^n$

$$(f_0 + g \circ F)^*(\bar{v}) \leq \text{cl} \rho(\bar{v}), \quad (3.5)$$

and for any $(\bar{v}, \bar{x}, \bar{y}) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m$

$$(f_0 + g \circ F)(\bar{x}) + (f_0 + g \circ F)^*(\bar{v}) = \langle \bar{v}, \bar{x} \rangle \quad (3.6a)$$

$$(f_0 + g \circ F)^*(\bar{v}) = (f_0 + \langle \bar{y}, F \rangle)^*(\bar{v}) + g^*(\bar{y}) \quad (3.6b)$$

is equivalent to

$$g(F(\bar{x})) + g^*(\bar{y}) = \langle \bar{y}, F(\bar{x}) \rangle \quad (3.7a)$$

$$(f_0 + \langle \bar{y}, F \rangle)(\bar{x}) + (f_0 + \langle \bar{y}, F \rangle)^*(\bar{v}) = \langle \bar{v}, \bar{x} \rangle. \quad (3.7b)$$

Finally, let $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$ be given by (3.1a) and

$$\begin{aligned} U &:= \left\{ u \in \mathbb{R}^m : \inf_{x \in \mathbb{R}^n} f(x, u) < +\infty \right\} \\ &= \text{dom}(g) - F(\text{dom}(f_0) \cap \text{dom}(F)). \end{aligned} \quad (3.8)$$

Assuming that $f \in \Gamma_0(\mathbb{R}^n \times \mathbb{R}^m)$ we have:

1. If $0 \in U$, then $\rho \in \Gamma(\mathbb{R}^n)$.
2. If $\rho \in \Gamma(\mathbb{R}^n)$ (e.g. if $0 \in U$), then (3.5) holds at equality.
3. If $0 \in \text{rint}(U)$ or $0 \in U$ and f is PWLQ, then $\rho \in \Gamma_0(\mathbb{R}^n)$, and the infimum in (3.3) is a minimum (possibly with $\rho(\bar{v}) = (f_0 + \langle y, F \rangle)^*(\bar{v}) + g^*(y) = +\infty$ for all $y \in \mathbb{R}^m$).

Proof First note that by using the characterization of f^* from Lemma 3.1 we have $\rho(\bar{v}) = q_0(\bar{v})$ and $(f_0 + g \circ F)^*(\bar{v}) = -p_{\bar{v}}(0)$ for all $\bar{v} \in \mathbb{R}^n$ where $q_{\bar{u}}$ and $p_{\bar{v}}$ are defined in (2.4). Then convexity of ρ and (3.5) follows from Theorem 2.1 (specifically (2.6)).

The characterizations from Lemma 3.1 also imply that (2.7) is equivalent to (3.6) and that (2.8) is equivalent to (3.7) as follows.

For the equivalence between (2.7) and (3.6) note that $\bar{x} \in P(0, \bar{v})$ is equivalent to (3.6a), $\bar{y} \in Q(0, \bar{v})$ is equivalent to $\rho(\bar{v}) = (f_0 + \langle \bar{y}, F \rangle)^*(\bar{v}) + g^*(\bar{y})$, and $p_{\bar{v}}(0) = -q_0(\bar{v})$ is equivalent to $(f_0 + g \circ F)^*(\bar{v}) = \rho(\bar{v})$. The equivalence then follows because (3.3) and (3.5) always imply

$$(f_0 + g \circ F)^*(\bar{v}) \leq \rho(\bar{v}) \leq (f_0 + \langle \bar{y}, F \rangle)^*(\bar{v}) + g^*(\bar{y}).$$

The equivalence between (2.8) and (3.7) follows by noting that under the characterization from Lemma 3.1, (2.8) becomes

$$\begin{aligned} \langle \bar{v}, \bar{x} \rangle - (f_0 + g \circ F)(\bar{x}) &= g^*(\bar{y}) - f_0(\bar{x}) - \langle \bar{y}, F \rangle(\bar{x}) + \langle \bar{v}, \bar{x} \rangle \\ &= g^*(\bar{y}) + (f_0 + \langle \bar{y}, F \rangle)^*(\bar{v}). \end{aligned}$$

The equivalence between (3.6) and (3.7) then follows from the corresponding equivalence between (2.7) and (2.8) in Theorem 2.1.

The remaining results follow from Corollary 2.1 by noting that $U = \text{dom}(p_0)$.

□

Next, in Section 3.1 we show how the equivalence between (3.6) and (3.7) can be used to deduce a chain rule for the sub-differential of $g \circ F$. After that in Section 3.2 how standard results on the sum of two functions can be used to refine both results (e.g. to write $(f_0 + \langle y, F \rangle)^*$ in (3.3) as a function of f_0^* and $(\langle y, F \rangle)^*$).

3.1. Sub-differential chain rule Combining the equivalence between (3.6) and (3.7) in Theorem 3.1 with Fenchel's inequality (e.g. [15, Proposition 11.3]) we also get the sub-differential chain rule for convex composite functions given in Corollary 3.1 below.

COROLLARY 3.1. *For $f_0 : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$, $F : \mathbb{R}^n \rightarrow \mathbb{R}^m \cup \{+\infty\}$ and $g : \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$, let $\langle y, F \rangle : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ given by (3.1b), $f_0 + g \circ F : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be given by (3.4), $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$ be given by (3.1a), and $U \subseteq \mathbb{R}^m$ be given by (3.8).*

If $f \in \Gamma(\mathbb{R}^n \times \mathbb{R}^m)$, $f(x, u)$ is closed in u , and $\text{dom}(F) \neq \emptyset$, then $f_0 + g \circ F$ and g are convex, $f_0 + \langle \bar{y}, F \rangle$ is convex for any $\bar{y} \in \text{dom}(g^)$ and for all $\bar{x} \in \text{dom}(f_0 + g \circ F)$ we have*

$$\partial(f_0 + g \circ F)(\bar{x}) \supseteq \bigcup_{\bar{y} \in \partial g(F(\bar{x}))} \partial(f_0 + \langle \bar{y}, F \rangle)(\bar{x}) \quad (3.9)$$

In addition, (3.9) holds at equality if $f \in \Gamma_0(\mathbb{R}^n \times \mathbb{R}^m)$ and either $0 \in \text{rint}(U)$ or $0 \in U$ and f is PWLQ.

Proof Convexity of $f_0 + g \circ F$ is direct from the joint convexity of f and convexity of g follows from convexity of f and the assumption $\text{dom}(F) \neq \emptyset$. Convexity of $f_0 + \langle \bar{y}, F \rangle$ follows by noting that by Theorem 2.2 and Lemma 3.1 we have that if $f(x, u)$ is convex in (x, u) and closed in u , then $l(x, y) = f_0(x) + \langle y, F(x) \rangle - g^*(y)$ is convex in x .

Now by Fenchel's inequality (e.g. [15, Proposition 11.3]) we have that

- (3.6a) is equivalent to $\bar{v} \in \partial(f_0 + g \circ F)(\bar{x})$,
- (3.7a) is equivalent to $\bar{y} \in \partial g(F(\bar{x}))$, and
- (3.7b) is equivalent to $\bar{v} \in \partial(f_0 + \langle \bar{y}, F \rangle)(\bar{x})$.

Then, (3.9) follows because (3.7a) and (3.7b) imply (3.6a) by Theorem 3.1.

Under the additional assumptions Theorem 3.1 implies that for any $\bar{x}, \bar{v} \in \mathbb{R}^n$ that satisfies (3.6a) there exists $\bar{y} \in \mathbb{R}^m$ that satisfies (3.6b). Equality in (3.9) then follows from the equivalence between (3.6) and (3.7). \square

3.2. Combining with standard additive results We can apply standard results for $(f_0 + \langle y, F \rangle)^*$ and $\partial(f_0 + \langle \bar{y}, F \rangle)(\bar{x})$ to further refine Theorem 3.1 and Corollary 3.1 in Proposition 3.1 below. Such standard results require $\langle \bar{y}, F \rangle \in \Gamma(\mathbb{R}^n)$, which the following lemma shows can be deduced from properties of the portion of f that does not depend on f_0 .

LEMMA 3.2. *For $f_0 : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$, $F : \mathbb{R}^n \rightarrow \mathbb{R}^m \cup \{+\infty\}$ and $g : \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$, let $\langle y, F \rangle : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ given by (3.1b), $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$ be given by (3.1a) and $h : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$ be given by*

$$h(x, u) := \begin{cases} g(F(x) + u) & x \in \text{dom}(F) \\ +\infty & \text{o.w.} \end{cases} \quad (3.10)$$

so that $f = f_0 + h$ (or equivalently so that $h = f$ with $f_0 \equiv 0$).

If $h \in \Gamma(\mathbb{R}^n \times \mathbb{R}^m)$ and $h(x, u)$ is closed in u , then $\langle \bar{y}, F \rangle \in \Gamma(\mathbb{R}^n)$ for all $\bar{y} \in \text{dom}(g^*)$.

Proof By Theorem 2.2 and Lemma 3.1 (applied to f with $f_0 \equiv 0$) we have that if $h(x, u)$ is convex in (x, u) and closed in u , then $l(x, y) = \langle y, F(x) \rangle - g^*(y)$ is convex in x . In addition, $h \in \Gamma(\mathbb{R}^n \times \mathbb{R}^m)$ implies $g \not\equiv +\infty$ and hence $g^*(y) > -\infty$ for all $y \in \mathbb{R}^m$. Then $\text{dom}(g^*) = \{y \in \mathbb{R}^m : g^*(y) \in \mathbb{R}\}$. We also have that $h \in \Gamma(\mathbb{R}^n \times \mathbb{R}^m)$ implies $\text{dom}(F) \neq \emptyset$. The result follows by noting that by its definition in (3.1b), $\langle \bar{y}, F \rangle$ is proper if and only if $\text{dom}(F) \neq \emptyset$. \square

PROPOSITION 3.1. *For $f_0 : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$, $F : \mathbb{R}^n \rightarrow \mathbb{R}^m \cup \{+\infty\}$ and $g : \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$, let $\langle y, F \rangle : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ given by (3.1b), $f_0 + g \circ F : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be given by (3.4), $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$ be given by (3.1a), $h : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$ be given by (3.10), $U \subseteq \mathbb{R}^m$ be given by (3.8), $\rho : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be given by (3.3) and $\tilde{\rho} : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be given by*

$$\tilde{\rho}(\bar{v}) := \inf_{y \in \mathbb{R}^m, w \in \mathbb{R}^n} f_0^*(w) + (\langle y, F \rangle)^*(\bar{v} - w) + g^*(y), \quad (3.11)$$

If $f_0 \in \Gamma(\mathbb{R}^n)$, $f, h \in \Gamma(\mathbb{R}^n \times \mathbb{R}^m)$, $h(x, u)$ is closed in u , then $f_0 + g \circ F$ is convex, and $f_0 + \langle \bar{y}, F \rangle, \langle \bar{y}, F \rangle \in \Gamma(\mathbb{R}^n)$ and for all $\bar{y} \in \text{dom}(g^*)$. In addition, for all $\bar{v} \in \mathbb{R}^n$ we have

$$(f_0 + g \circ F)^*(\bar{v}) \leq \rho(\bar{v}) \leq \tilde{\rho}(\bar{v}), \quad (3.12)$$

and for all $\bar{x} \in \text{dom}(f_0 + g \circ F)$ we have

$$\begin{aligned} \partial(f_0 + g \circ F)(\bar{x}) &\supseteq \bigcup_{\bar{y} \in \partial g(F(\bar{x}))} \partial(f_0 + \langle \bar{y}, F \rangle)(\bar{x}) \\ &\supseteq \partial f_0(\bar{x}) + \bigcup_{\bar{y} \in \partial g(F(\bar{x}))} \partial \langle \bar{y}, F \rangle(\bar{x}). \end{aligned} \quad (3.13)$$

In addition, equality holds throughout (3.12) and (3.13) if additionally $f \in \Gamma_0(\mathbb{R}^n \times \mathbb{R}^m)$,

$$\text{rint}(\text{dom}(f_0)) \cap \text{rint}(\text{dom}(F)) \neq \emptyset \quad (3.14)$$

and either $0 \in \text{rint}(U)$ or $0 \in U$ and f is PWLQ.

Proof Convexity of $f_0 + g \circ F$ is direct from $f \in \Gamma(\mathbb{R}^n \times \mathbb{R}^m)$. Lemma 3.2 implies $\langle \bar{y}, F \rangle \in \Gamma(\mathbb{R}^n)$ for all $\bar{y} \in \text{dom}(g^*)$. $f \in \Gamma(\mathbb{R}^n \times \mathbb{R}^m)$ also implies $\text{dom}(f_0) \cap \text{dom}(F) \neq \emptyset$. Then $f_0 \in \Gamma(\mathbb{R}^n)$ and $\langle \bar{y}, F \rangle \in \Gamma(\mathbb{R}^n)$ for all $\bar{y} \in \text{dom}(g^*)$ implies $f_0 + \langle \bar{y}, F \rangle \in \Gamma(\mathbb{R}^n)$ for all $\bar{y} \in \text{dom}(g^*)$.

The first inequality in (3.12) and its equality under the additional assumptions follows from Theorem 3.1. The second inequality in (3.12) and its equality under the additional assumptions follows from [12, Theorem 16.4], which is applicable because $f_0, \langle \bar{y}, F \rangle \in \Gamma(\mathbb{R}^n)$ for all $\bar{y} \in \text{dom}(g^*)$.

The first inclusion in (3.13) and its equality under the additional assumptions follows from Corollary 3.1. The second inclusion in (3.13) and its equality under the additional assumptions follows from [12, Theorem 23.8], which is also applicable because $f_0, \langle \bar{y}, F \rangle \in \Gamma(\mathbb{R}^n)$ for all $\bar{y} \in \text{dom}(g^*)$ and $\bar{y} \in \partial g(F(\bar{x}))$ implies $\bar{y} \in \text{dom}(g^*)$. \square

4. Connection of Theorem 3.1 to K -convexity A standard sufficient condition for convexity of $g \circ F$ is component-wise convexity of F and component-wise monotonicity of g (e.g. [9]). Such condition can be further generalized through the following notions of convexity and monotonicity with regards to a closed convex cone K [6, 8, 11].

DEFINITION 4.1. Given a closed convex cone $K \subseteq \mathbb{R}^m$ we define the ordering

$$x^1 \leq_K x^2 \iff x^2 - x^1 \in K \quad \forall x^1, x^2 \in \mathbb{R}^m.$$

We also extend the ordering to consider the infinite element $+\infty_\bullet$ (cf. Definition 3.1) as its *largest element* such that

$$x \leq_K +\infty_\bullet \quad \forall x \in \mathbb{R}^m.$$

DEFINITION 4.2. Let $K \subseteq \mathbb{R}^m$ be a closed convex cone and $+\infty_\bullet$ be a *largest element* with respect to K (cf. Definition 4.1).

We say $F : \mathbb{R}^n \rightarrow \mathbb{R}^m \cup \{+\infty_\bullet\}$ is K -convex if and only if

$$K\text{-epi}(F) := \{(x, z) \in \text{dom}(F) \times \mathbb{R}^m : F(x) \leq_K z\}$$

is convex.

We say $g : \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$ is K -increasing if and only if

$$g(w) \leq g(w + k) \quad \forall k \in K, w \in \mathbb{R}^m,$$

and is K -increasing restricted to $\text{rge}(F)$ if and only if

$$g(w) \leq g(w + k) \quad \forall k \in K, w \in \text{rge}(F).$$

Finally, we let

$$\begin{aligned}\mathcal{K}_F &:= \{K \subseteq \mathbb{R}^m : K \text{ is a closed convex cone and } F \text{ is } K\text{-convex}\} \\ \mathcal{K}_g &:= \{K \subseteq \mathbb{R}^m : K \text{ is a closed convex cone and } g \text{ is } K\text{-increasing}\} \\ \mathcal{K}_{g,\text{rng}(F)} &:= \left\{ K \subseteq \mathbb{R}^m : \begin{array}{l} K \text{ is a closed convex cone and} \\ g \text{ is } K\text{-increasing restricted to } \text{rge}(F) \end{array} \right\}\end{aligned}$$

PROPOSITION 4.1. *Let $g : \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$, $F : \mathbb{R}^n \rightarrow \mathbb{R}^m \cup \{+\infty\}$, $g \circ F : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be given by*

$$g \circ F(x) := \begin{cases} g(F(x)) & x \in \text{dom}(F) \\ +\infty & \text{o.w.} \end{cases}, \quad (4.1)$$

and $f_0 : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be such that $f_0 \equiv 0$ so that $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$ given by (3.1a) becomes

$$f(x, u) := \begin{cases} g(F(x) + u) & x \in \text{dom}(F) \\ +\infty & \text{o.w.} \end{cases} \quad (4.2)$$

and $g \circ F(x) = f(x, 0)$ for all $x \in \mathbb{R}^n$.

If $g \in \Gamma(\mathbb{R}^m)$, then

$$\mathcal{K}_F \cap \mathcal{K}_{g,\text{rng}(F)} \neq \emptyset \Rightarrow g \circ F \text{ is convex} \quad (4.3a)$$

$$\mathcal{K}_F \cap \mathcal{K}_g \neq \emptyset \Rightarrow f \text{ is convex} \quad (4.3b)$$

If $g \in \Gamma_0(\mathbb{R}^m)$ and $\text{dom}(g) \cap \text{rge}(F) \neq \emptyset$, then (4.3b) is an equivalence.

→ See page 26 for the proof.

As illustrated in [8, Example 58], convexity of $g \circ F$ can be deduced through $\mathcal{K}_F \cap \mathcal{K}_{g,\text{rng}(F)} \neq \emptyset$ for cases in which $\mathcal{K}_F \cap \mathcal{K}_g \neq \emptyset$ does not hold. This is not surprising given that the latter condition implies convexity of f , which is stronger than just convexity of $g \circ F$. Based on condition $\mathcal{K}_F \cap \mathcal{K}_{g,\text{rng}(F)} \neq \emptyset$, [6] also studies characterizations of $(g \circ F)^*$ as in Theorem 3.1. The first version of this characterization is [6, Theorem 4] that focuses on inequality (3.5) from Theorem 3.1 when $f_0 \equiv 0$. In this setting, we can restate [6, Theorem 4] as follows.

THEOREM 4.1 ([6, Theorem 4]). *Let $K \subseteq \mathbb{R}^m$ be a closed convex cone, $g : \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$, $F : \mathbb{R}^n \rightarrow \mathbb{R}^m \cup \{+\infty\}$, $\langle y, F \rangle : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be given by (3.1b), $g \circ F : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be given by (4.1) and $\eta : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be given by*

$$\eta(v) := \inf_{y \in -K^\circ} (\langle y, F \rangle)^*(v) + g^*(y), \quad (4.4)$$

If $g \in \Gamma(\mathbb{R}^m)$, and $K \in \mathcal{K}_F \cap \mathcal{K}_{g, \text{rng}(F)}$, then for any $\bar{v} \in \mathbb{R}^n$

$$(g \circ F)^*(\bar{v}) \leq \text{cl } \eta(\bar{v}). \quad (4.5)$$

If $g \in \Gamma_0(\mathbb{R}^m)$, $K \in \mathcal{K}_F \cap \mathcal{K}_g$, $\text{dom}(g) \cap \text{rge}(F) \neq \emptyset$, and $\langle y, F \rangle \in \Gamma_0(\mathbb{R}^n)$ for all $y \in -K^\circ$, then (4.5) holds at equality.

Finally, if $g \in \Gamma(\mathbb{R}^m)$, $K \in \mathcal{K}_F \cap \mathcal{K}_{g, \text{rng}(F)}$ and

$$\text{rint}(\text{dom}(g) - K) \cap F(\text{rint}(\text{dom}(F))) \neq \emptyset, \quad (4.6)$$

then $\eta \in \Gamma_0(\mathbb{R}^n)$ and the infimum in (4.4) is a minimum (possibly with $\eta(v) = (\langle y, F \rangle)^(v) + g^*(y) = +\infty$ for all $y \in -K^\circ$).*

A chain rule for the sub-differential of $g \circ F$ is also presented in [6, Corollary 2]. This result is identical to Corollary 3.1 except for the necessary conditions for (3.9) to hold (i) as is and (ii) as an equality. These conditions are the same as those (i) for (4.5) to hold and (ii) for (4.5) to hold at equality, η being closed and then infimum in (4.4) being attained. Hence, a comparison between [6, Corollary 2] and Corollary 3.1 follows directly from the comparison between 4.1 and Theorem 3.1 in Section 4.1.

Another corollary of Theorem 4.1 considered in [6] is [6, Corollary 3] which extends the convex conjugate and sub-differential formulas for the case $f_0 \not\equiv 0$. We compare [6, Corollary 3] with Proposition 3.1 in Section 4.3.

4.1. Comparison between Theorems 3.1 and 4.1 Under different requirements, both $\eta : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ from Theorem 4.1 and $\rho : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ from Theorem 3.1 (with $f_0 \equiv 0$) yield upper bounds on $(g \circ F)^*$ with ρ being a potentially tighter bound:

$$(g \circ F)^*(\bar{v}) \leq \text{cl } \rho(\bar{v}) \leq \text{cl } \eta(\bar{v}) \quad \forall \bar{v} \in \mathbb{R}^n. \quad (4.7)$$

It is worth noting that the only reason that ρ is potentially stronger than η is that the definition of η in (4.4) restricts the infimum over values $y \in -K^\circ$. We will further study this in Section 4.2.

To compare the requirements of Theorems 3.1 and 4.1 we will use the following lemma.

LEMMA 4.1. *Let $K \subseteq \mathbb{R}^m$ be a closed convex cone, $g : \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$, $F : \mathbb{R}^n \rightarrow \mathbb{R}^m \cup \{+\infty\}$, $\langle y, F \rangle : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be given by (3.1b) and $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$ given by (4.2).*

If $g \in \Gamma_0(\mathbb{R}^m)$, $K \in \mathcal{K}_F \cap \mathcal{K}_g$, $\text{dom}(g) \cap \text{rge}(F) \neq \emptyset$, and $\langle y, F \rangle \in \Gamma_0(\mathbb{R}^n)$ for all $y \in -K^\circ$, then $f \in \Gamma_0(\mathbb{R}^n \times \mathbb{R}^m)$.

→ See page 26 for the proof.

With regards to $\text{cl } \rho$ and $\text{cl } \eta$ providing upper bounds on $(g \circ F)^*$ we see that Theorem 3.1 does not require anything for it to hold, while Theorem 4.1 requires the existence of a closed convex cone $K \in \mathcal{K}_F \cap \mathcal{K}_{g, \text{rng}(F)}$, which is used to define η . By Proposition 4.1 these conditions imply convexity of $g \circ F$ which may be considered a natural setting for studying $(g \circ F)^*$. Nonetheless, $(g \circ F)^*$ is convex even if $g \circ F$ is not and Theorem 3.1 provides ρ as a convex upper bound of $(g \circ F)^*$ even in such non-convex setting.

With regards to $\text{cl } \rho$ and $\text{cl } \eta$ being equal to $(g \circ F)^*$, by Lemma 4.1 we have that the requirements of Theorem 4.1 imply that $f \in \Gamma_0(\mathbb{R}^n \times \mathbb{R}^m)$ and $0 \in \text{dom}(g) - \text{rng}(F)$, which are essentially the requirements of Theorem 3.1 ($0 \in \text{dom}(g) - \text{rng}(F)$ implies ρ being proper, which is the real requirement of Theorem 3.1). Combining this with Proposition 4.1 we see that the requirements of Theorem 3.1 for $(g \circ F)^* = \text{cl } \rho$ imply the existence of a closed convex cone $K \in \mathcal{K}_F \cap \mathcal{K}_g$. The requirement $K \in \mathcal{K}_F \cap \mathcal{K}_g$ also appears in Theorem 4.1 for $(g \circ F)^* = \text{cl } \eta$. Hence the additional flexibility of using $\mathcal{K}_{g, \text{rng}(F)}$ instead of \mathcal{K}_g does not seem to provide an advantage when approximating $(g \circ F)^*$ (getting an upper bound does not require it and it is not sufficient to guarantee a tight upper bound).

With regards to η being closed and that the infimum in its definition being attained, we do have that the weaker requirement $K \in \mathcal{K}_F \cap \mathcal{K}_{g, \text{rng}(F)}$ from Theorem 4.1 can provide an advantage over the more restrictive $K \in \mathcal{K}_F \cap \mathcal{K}_g$ required by Theorem 3.1 for the same properties holding for ρ . However, it could be argued that such properties for η are less interesting when $\text{cl } \eta$ is a strict upper bound on $(g \circ F)^*$. For this reason we first restrict our attention to the case where Theorem 3.1 and 4.1 imply $(g \circ F)^* = \text{cl } \rho = \text{cl } \eta$ and then in Section 4.2 study how the existence of $K \in \mathcal{K}_F \cap \mathcal{K}_{g, \text{rng}(F)} \setminus \mathcal{K}_g$ can still be useful in this context. For simplicity we consider the following

joint condition for $(g \circ F)^* = \text{cl } \rho = \text{cl } \eta$. First we assume

$$g \in \Gamma_0(\mathbb{R}^m), \quad (4.8a)$$

$$\text{dom}(g) \cap \text{rge}(F) \neq \emptyset, \quad (4.8b)$$

$$f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \overline{\mathbb{R}} \quad \text{is closed,} \quad (4.8c)$$

$$\langle y, F \rangle : \mathbb{R}^n \rightarrow \overline{\mathbb{R}} \quad \text{is proper and closed } \forall y \in \mathbb{R}^m. \quad (4.8d)$$

Now, by [6, Theorem 3] we have that $K \in \mathcal{K}_F$ if and only if $\langle y, F \rangle : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ is convex for all $y \in -K^\circ$. Then, by Proposition 4.1 a joint condition for $(g \circ F)^* = \text{cl } \rho = \text{cl } \eta$ is given by (4.8) plus either of the following equivalent conditions

$$\mathcal{K}_F \cap \mathcal{K}_g \neq \emptyset, \quad (4.9a)$$

$$f \text{ is convex.} \quad (4.9b)$$

More precisely, for $(g \circ F)^* = \text{cl } \eta$, the condition is that we use $K \in \mathcal{K}_F \cap \mathcal{K}_g$ to define η through (4.4). To compare the closure and infimum-attainment conditions from Theorem 3.1 and 4.1 in this context we will use the following lemma.

LEMMA 4.2. *If (4.8) and (4.9a) (or equivalently (4.9b)) holds for g and F , then the general condition $0 \in \text{rint}(U)$ from Theorem 3.1 for $f_0 \equiv 0$ is equivalent to*

$$\exists K \in \mathcal{K}_F \cap \mathcal{K}_g \quad s.t. \quad \text{rint}(\text{dom}(g)) \cap \text{rint}(F(\text{dom}(F)) + K) \neq \emptyset, \quad (4.10)$$

the weaker condition for piecewise linear-quadratic functions $0 \in U$ from Theorem 3.1 for $f_0 \equiv 0$ is equivalent to

$$\exists K \in \mathcal{K}_F \cap \mathcal{K}_g \quad s.t. \quad (\text{dom}(g) - K) \cap (F(\text{dom}(F))) \neq \emptyset, \quad (4.11)$$

and condition (4.6) from Theorem 4.1 is equivalent to

$$K \in \mathcal{K}_F \cap \mathcal{K}_g \quad \text{and} \quad \text{rint}(\text{dom}(g) - K) \cap F(\text{rint}(\text{dom}(F))) \neq \emptyset, \quad (4.12)$$

→ See page 27 for the proof.

As noted in [6, pages 1335–1336], condition (4.10) matches the Attouch–Brézis-type constraint qualification conditions from [7] and is in general incomparable with (4.12) (each being weaker than the other in some cases). In contrast, condition (4.11) is weaker than the two others as expected from it only being sufficient when f is a piecewise linear-quadratic function.

For all three conditions, we see that it will be useful to understand properties of the largest cone $K \in \mathcal{K}_F \cap \mathcal{K}_g$ for which we can rely on the following result.

LEMMA 4.3 ([8, Lemma 10, Proposition 24 and 55]). *Let $F : \mathbb{R}^n \rightarrow \mathbb{R}^m \cup \{+\infty\}$ be such that $\text{dom}(F)$ is non-empty and convex. Then there exist a non-empty, closed and convex cone $K_F \subseteq \mathbb{R}^m$ such that F is K -convex for a closed convex cone $K \subseteq \mathbb{R}^m$ if and only if $K_F \subseteq K$.*

Let $g \in \Gamma_0(\mathbb{R}^m)$. Then g is K -increasing for a closed convex cone $K \subseteq \mathbb{R}^m$ if and only if $K \subseteq -\text{hzn}(g)$, where

$$\text{hzn}(g) := \{d \in \mathbb{R}^m : g(x+d) \leq g(x) \quad \forall x \in \text{dom}(g)\}$$

is the horizon cone of g .

In particular, we have that

$$\mathcal{K}_F \cap \mathcal{K}_g = \left\{ K \subseteq \mathbb{R}^m : \begin{array}{l} K \text{ is a closed convex cone and} \\ K_F \subseteq K \subseteq -\text{hzn}(g) \end{array} \right\}.$$

for any closed convex cone $K \subseteq \mathbb{R}^m$ we have that F is K -convex and g is K -increasing if and only if $K_F \subseteq K \subseteq -\text{hzn}(g)$.

4.2. Using K -convexity to satisfy necessary conditions We illustrate the potential value of a cone $K \in \mathcal{K}_{F,\text{rng}(F)} \setminus \mathcal{K}_g$ through the following two examples.

EXAMPLE 4.1 ([6, VARIANT OF EXAMPLE 1]). Let $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be given by $F(x) := (0, x_2)$ and $g : \mathbb{R}^2 \rightarrow \overline{\mathbb{R}}$ be given by

$$g(w) := \begin{cases} \frac{1}{3}w_1^3 & w_1 \geq 0 \\ +\infty & \text{o.w.} \end{cases}.$$

We have that $g \circ F \equiv 0$, $(g \circ F)^* = \delta_{\{(0,0)\}}$, $\text{dom}(g) = \mathbb{R}_+ \times \mathbb{R}$, $\text{dom}(F) = \mathbb{R}^2$, $\text{rge}(F) = \{0\} \times \mathbb{R}$ and $f(x, u) = g(u)$ for all $x \in \mathbb{R}^2$ and $u \in \mathbb{R}^2$. Then $f \in \Gamma_0(\mathbb{R}^2 \times \mathbb{R}^2)$ and $0 \in U = \text{dom}(g) - \text{rng}(F) = \mathbb{R}_+ \times \mathbb{R}$, so by Theorem 3.1 we have $(g \circ F)^* = \text{cl}(\rho)$. Indeed,

$$g^*(v) := \begin{cases} \frac{2}{3}v_1^{3/2} & v_1 \geq 0, v_2 = 0 \\ 0 & v_1 < 0, v_2 = 0 \\ +\infty & \text{o.w.} \end{cases}$$

and $(\langle y, F \rangle)^* = \delta_{\{0\} \times \{y_2\}}$, so

$$\rho(v) = \inf_{y \in \mathbb{R}^2} \delta_{\{0\} \times \{y_2\}}(v) + g^*(y) = \delta_{\{0\} \times \mathbb{R}}(v) + g^*(0, v_2) = \delta_{\{(0,0)\}}(v). \quad (4.13)$$

Hence, ρ is additionally closed and the infimum in its definition is attained. Unfortunately, f is not piecewise linear-quadratic and $0 \notin \text{rint}(U) = \text{rint}(\mathbb{R}_+ \times \mathbb{R})$ so condition (4.10) is not satisfied and hence Theorem 3.1 cannot be used to deduce these additional properties.

Now, we also have $K_F = \{(0,0)\} \subseteq -\text{hzn}(g) = \{0\} \times \mathbb{R}$, so by Lemma 4.3 the largest cone in $\mathcal{K}_F \cap \mathcal{K}_g$ is $K_1 := \{0\} \times \mathbb{R}$. As expected,

$$\text{rint}(\text{dom}(g)) \cap \text{rint}(F(\text{dom}(F)) + K) \neq \emptyset \quad (4.14)$$

does not hold for $K = K_1$ as (4.10) does not hold. However, (4.14) does hold for $K = K_2 := \mathbb{R}_+ \times \{0\} \in \mathcal{K}_F \cap \mathcal{K}_{g,\text{rng}(F)} \setminus \mathcal{K}_g$.

EXAMPLE 4.2 ([8, EXAMPLE 58]). Let $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be given by $F(x) := \left(\frac{1}{2}x_1^2, x_2\right)$ and $g : \mathbb{R}^2 \rightarrow \overline{\mathbb{R}}$ be given by $g(w) = |w_1|$.

We have that $g \circ F(x) = \frac{1}{2}w_1^2$, $(g \circ F)^*(v) = \frac{1}{2}v_1^2 + \delta_{\{0\}}(v_2)$, $g^* = \delta_{[-1,1] \times \{0\}}$,

$$(\langle y, F \rangle)^*(v) = \begin{cases} \frac{1}{2y_1}v_1^2 & y_1 > 0, y_2 = v_2 \\ +\infty & \text{o.w.} \end{cases},$$

so

$$\rho(v) = \inf_{\substack{y_1 > 0 \\ y_2 \in \mathbb{R}}} \frac{1}{2y_1}v_1^2 + \delta_{\{y_2\}}(v_2) + g^*(y) = \frac{1}{2}v_1^2 + g^*(1, v_2) = \frac{1}{2}v_1^2 + \delta_{\{0\}}(v_2).$$

Hence, $(g \circ F)^* = \rho$, $\rho \in \Gamma_0(\mathbb{R}^2)$ and the infimum in the definition of ρ is always attained. Unfortunately, $f(x, u) = g(F(x) + u) = \left|\frac{1}{2}x_1^2 + u_1\right|$ is not convex and hence Theorem 3.1 cannot be used to deduce these properties.

Now, we can also check that $K_F = \mathbb{R}_+ \times \{0\}$ and $\text{hzn}(g) = \{0\} \times \mathbb{R}$ so by Lemma 4.3 we have that $\mathcal{K}_F \cap \mathcal{K}_g$ is empty. However, we also have $K_F \in \mathcal{K}_F \cap \mathcal{K}_{g,\text{rng}(F)} \setminus \mathcal{K}_g$.

In both Examples 4.1 and 4.2 we have a cone $K \in \mathcal{K}_F \cap \mathcal{K}_{g,\text{rng}(F)} \setminus \mathcal{K}_g$, which we would like to also be in $\mathcal{K}_F \cap \mathcal{K}_g$. Now, the motivation for the definition of $\mathcal{K}_{g,\text{rng}(F)}$ is that the behavior of g outside of $\text{rng}(F)$ does not affect $g \circ F$. This suggests that we may be able to modify g to achieve $\mathcal{K}_{g,\text{rng}(F)} = \mathcal{K}_g$ without changing $g \circ F$. Such modification is precisely what is provided by the following observation from [11, page 253]¹

¹ Note that the definitions of K -convexity, K -epigraph, etc. in this reference are reversed compared to those in [6, 8]. For instance, in [6, page 1328] it is noted that for $h : \mathbb{R} \rightarrow \overline{\mathbb{R}}$ convexity corresponds to K -convexity with $K = \mathbb{R}_+$, while in [11, page 236] it is noted that for their version of it corresponds to K -convexity with $K = \mathbb{R}_-$.

PROPOSITION 4.2 ([11]). *Let $g : \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$, $F : \mathbb{R}^n \rightarrow \mathbb{R}^m \cup \{+\infty\}$ and $K \in \mathcal{K}_F \cap \mathcal{K}_{g, \text{rng}(F)}$. Then $g_K : \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$ given by*

$$g_K(w) := g \square \delta_{-K} = \inf_{k \in K} g(w + k) \quad (4.15)$$

is such that

$$g(w) = g_K(w) \quad \forall w \in \text{rge}(F). \quad (4.16)$$

In addition, $K \in \mathcal{K}_F \cap \mathcal{K}_{g_K}$.

Replacing g with g_K for $K \in \mathcal{K}_F \cap \mathcal{K}_{g, \text{rng}(F)}$ then has the potential to repair the issues in Examples 4.1 and 4.2. However, while replacing g with g_K does not change $g \circ F$, it does change g^* and hence ρ . Fortunately, using standard results (e.g. [15, Theorem 11.23a]) we get the following simple characterization of g_K^* as a function of g^* .

LEMMA 4.4. *Let $K \subseteq \mathbb{R}^m$ be a closed convex cone, $g : \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$ be a proper function and $g_K : \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$ given by (4.15). Then $g_K^* = g^* + \delta_{-K^\circ}$.*

Using Lemma 4.4 we see that the effect of replacing g with g_K in definition (3.3) of ρ is replacing $y \in \mathbb{R}^m$ by $y \in -K^\circ$, which results precisely in definition (4.4) of η and explains the difference between the definitions of ρ and η^2 .

Using Proposition 4.2 and Lemma 4.4 we can resolve the issues in Examples 4.1 and 4.2 as follows.

EXAMPLE 4.3 (EXAMPLE 4.1 CONTINUED). Using Proposition 4.2 for $K_2 := \mathbb{R}_+ \times \{0\} \in \mathcal{K}_F \cap \mathcal{K}_{g, \text{rng}(F)}$ we get $g_{K_2} : \mathbb{R}^2 \rightarrow \overline{\mathbb{R}}$ given by

$$g_{K_2}(w) := \begin{cases} \frac{1}{3}w_1^3 & w_1 \geq 0 \\ 0 & \text{o.w.} \end{cases},$$

for which $g_{K_2} \circ F = g \circ F$ and $\text{dom}(g_{K_2}) = \mathbb{R}^2$. In addition, for $f_{K_2} : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \overline{\mathbb{R}}$ given by $f_{K_2}(x, u) = g_{K_2}(F(x) + u) = g_{K_2}(u)$ we have $f_{K_2} \in \Gamma_0(\mathbb{R}^2 \times \mathbb{R}^2)$, and $0 \in \text{rint}(U_{K_2}) = \text{rint}(\text{dom}(g_{K_2}) - \text{rng}(F)) = \mathbb{R}^2$. Hence, we can now apply Theorem 3.1 to conclude that have $(g \circ F)^* = (g_{K_2} \circ F)^* = \rho_{K_2}$ where $\rho_{K_2} : \mathbb{R}^2 \rightarrow \overline{\mathbb{R}}$ is given by

$$\rho_{K_2}(v) = \inf_{y \in \mathbb{R}^2} \delta_{\{0\} \times \{y_2\}}(v) + g_{K_2}^*(y) = \delta_{\{0\} \times \mathbb{R}}(v) + g_{K_2}^*(0, v_2) = \delta_{\{(0, 0)\}}(v),$$

² cf. [6, Lemma 4] and the ψ^* in [6, Lemma 5] for which we have $f = \psi^*$.

where the last equation follows from Lemma 4.4:

$$g_{K_2}^*(v) = g^*(v) + \delta_{-K_2^\circ}(v) = g^*(v) + \delta_{\mathbb{R}_+ \times \mathbb{R}}(v) = \begin{cases} \frac{2}{3}v_1^{3/2} & v_1 \geq 0, v_2 = 0 \\ +\infty & \text{o.w.} \end{cases}.$$

EXAMPLE 4.4 (EXAMPLE 4.2 CONTINUED). Using Proposition 4.2 for $K_F = \mathbb{R}_+ \times \{0\} \in \mathcal{K}_F \cap \mathcal{K}_{g, \text{rng}(F)}$ we get $g_{K_F} : \mathbb{R}^2 \rightarrow \overline{\mathbb{R}}$ given by $g_{K_F}(w) := \max\{w_1, 0\}$ for which $g_{K_F} \circ F = g \circ F$ and $\text{dom}(g_{K_F}) = \mathbb{R}^2$. In addition, for $f_{K_F} : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \overline{\mathbb{R}}$ given by $f_{K_F}(x, u) = g_{K_F}(F(x) + u) = \max\{\frac{1}{2}x_1^2 + u_1, 0\}$ we have $f_{K_F} \in \Gamma_0(\mathbb{R}^2 \times \mathbb{R}^2)$ and $0 \in \text{rint}(U_{K_F}) = \text{rint}(\text{dom}(g_{K_F}) - \text{rng}(F)) = \mathbb{R}^2$. Hence, we can now apply Theorem 3.1 to conclude that have $(g \circ F)^* = (g_{K_F} \circ F)^* = \rho_{K_F}$ where $\rho_{K_F} : \mathbb{R}^2 \rightarrow \overline{\mathbb{R}}$ is given by

$$\rho_{K_F}(v) = \inf_{\substack{y_1 > 0 \\ y_2 \in \mathbb{R}}} \frac{1}{2}y_1^2 + \delta_{\{y_2\}}(v_2) + g_{K_F}^*(y) = \frac{1}{2}v_1^2 + g_{K_F}^*(1, v_2) = \frac{1}{2}v_1^2 + \delta_{\{0\}}(v_2).$$

where the last equation follows from Lemma 4.4:

$$g_{K_F}^* = g^* + \delta_{-K_F^\circ} = \delta_{[-1, 1] \times \{0\}} + \delta_{\mathbb{R}_+ \times \{0\}} = \delta_{[0, 1] \times \{0\}}.$$

We end this section by noting that while g_K preserves convexity of g it does not necessarily preserve g being proper or closed. Hence, repairing a bad choice of g to describe $g \circ F$ may require additional steps such as first restricting g to $\text{rng}(F)$ (if this set is convex) or $\text{rng}(F) + K$ (which is convex if F is K -convex). This is illustrated in the following example.

EXAMPLE 4.5. Let $K = \mathbb{R}_+ \times \{0\}$ and $g : \mathbb{R}^2 \rightarrow \overline{\mathbb{R}}$ be given by

$$g(w) := \begin{cases} -\sqrt{w_1 w_2} & w_1 \geq 0, w_2 \geq 0 \\ +\infty & \text{o.w.} \end{cases}.$$

Assume we want to compose g with a K -convex function $F : \mathbb{R} \rightarrow \mathbb{R}^2 \cup \{+\infty\}$ for which $\text{rng}(F) = \mathbb{R}_+ \times \{0\}$. Using Proposition 4.2 for $K \in \mathcal{K}_F \cap \mathcal{K}_{g, \text{rng}(F)} \setminus \mathcal{K}_g$ we get $g_K : \mathbb{R}^2 \rightarrow \overline{\mathbb{R}}$ given by

$$g_K(w) = \begin{cases} -\infty & w_2 > 0 \\ 0 & w_2 = 0 \\ +\infty & \text{o.w.} \end{cases},$$

which is not proper. However, if we first restrict g to $\text{rge}(F) = \text{rge}(F) + K = \mathbb{R}_+ \times \{0\}$ we get $(g + \delta_{\text{rge}(F)})_K : \mathbb{R}^2 \rightarrow \overline{\mathbb{R}}$ given by

$$(g + \delta_{\text{rge}(F)})_K(w) = \begin{cases} 0 & w_2 = 0 \\ +\infty & \text{o.w.} \end{cases},$$

for which $(g + \delta_{\text{rge}(F)})_K \in \Gamma_0(\mathbb{R}^2)$.

4.3. Additive composite functions Using standard results on the conjugate of the sum of two functions (e.g. [12, Theorem 16.4]), [6] also extends Theorem 4.1 to the case $f_0 \not\equiv 0$ through [6, Corollary 3], which we can restate as follows.

COROLLARY 4.1 ([6, Corollary 3]). *Let $K \subseteq \mathbb{R}^m$ be a closed convex cone, $f_0 : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$, $g : \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$, $F : \mathbb{R}^n \rightarrow \mathbb{R}^m \cup \{+\infty_\bullet\}$, $\langle y, F \rangle : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be given by (3.1b), and $f_0 + g \circ F : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be given by (3.4).*

If $g \in \Gamma_0(\mathbb{R}^m)$, $y_0 f_0 + \langle y, F \rangle \in \Gamma_0(\mathbb{R}^n)$ for all $(y_0, y) \in \mathbb{R}_+ \times (-K^\circ)$, $K \in \mathcal{K}_F \cap \mathcal{K}_g$, and

$$\text{rint}(\text{dom}(g) - K) \cap F(\text{rint}(\text{dom}(f_0)) \cap \text{rint}(\text{dom}(F))) \neq \emptyset, \quad (4.17)$$

then, for any $\bar{v} \in \mathbb{R}^n$ we have

$$(f_0 + g \circ F)^*(\bar{v}) = \min_{y \in -K^\circ, w \in \mathbb{R}^n} f_0^*(w) + (\langle y, F \rangle)^*(\bar{v} - w) + g^*(y). \quad (4.18)$$

From Proposition 3.1 we can get the following analogous corollary

COROLLARY 4.2. *For $f_0 : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$, $F : \mathbb{R}^n \rightarrow \mathbb{R}^m \cup \{+\infty_\bullet\}$ and $g : \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$, let $\langle y, F \rangle : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ given by (3.1b), $f_0 + g \circ F : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be given by (3.4) and $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$ be given by (3.1a).*

If $f_0 \in \Gamma(\mathbb{R}^n)$, $h \in \Gamma(\mathbb{R}^n \times \mathbb{R}^m)$, $f \in \Gamma_0(\mathbb{R}^n \times \mathbb{R}^m)$,

$$\text{rint}(\text{dom}(f_0)) \cap \text{rint}(\text{dom}(F)) \neq \emptyset \quad (4.19)$$

and there exists $K \in \mathcal{K}_F \cap \mathcal{K}_g$ such that either

$$\text{rint}(\text{dom}(g)) \cap \text{rint}(F(\text{dom}(f_0) \cap \text{dom}(F)) + K) \neq \emptyset, \quad (4.20)$$

or f is PWLQ, and

$$(\text{dom}(g) - K) \cap (F(\text{dom}(f_0) \cap \text{dom}(F))) \neq \emptyset, \quad (4.21)$$

then, for any $\bar{v} \in \mathbb{R}^n$ we have

$$(f_0 + g \circ F)^*(\bar{v}) = \min_{y \in \mathbb{R}^m, w \in \mathbb{R}^n} f_0^*(w) + (\langle y, F \rangle)^*(\bar{v} - w) + g^*(y), \quad (4.22)$$

Similarly to the case $f_0 \equiv 0$ we can see that PWLQ conditions (4.19) and (4.21) from Corollary 4.2 are weaker than condition (4.17) from Corollary 4.1, while general conditions (4.19) and (4.20) from Corollary 4.2 and (4.17) from Corollary 4.1 are incomparable.

Finally, note that for Corollary 4.2 we can relax (4.19) if we have additional information on f_0 and F . For instance, based on the comments after [10, Theorem E.2.3.2] we can relax (4.19) to $\text{dom}(f_0) \cap \text{dom}(F) \neq \emptyset$ if f_0 and $\langle \bar{y}, F \rangle$ are polyhedral for all $y \in \mathbb{R}^m$.

5. Omitted Proofs

5.1. Proofs from Section 3

LEMMA 3.1. Let $f_0 : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$, $F : \mathbb{R}^n \rightarrow \mathbb{R}^m \cup \{+\infty\}$ and $g : \mathbb{R}^m \rightarrow \bar{\mathbb{R}}$ let $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \bar{\mathbb{R}}$, $\langle y, F \rangle : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$ and $l : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \bar{\mathbb{R}}$ be given by

$$f(x, u) := \begin{cases} f_0(x) + g(F(x) + u) & x \in \text{dom}(F) \\ +\infty & \text{o.w.} \end{cases}, \quad (3.1a)$$

$$\langle y, F \rangle(x) := \begin{cases} \langle y, F(x) \rangle & x \in \text{dom}(F) \\ +\infty & \text{o.w.} \end{cases}, \quad (3.1b)$$

$$l(x, y) := -(f(x, \cdot))^*(y). \quad (3.1c)$$

Then

$$l(x, y) = f_0(x) + \langle y, F(x) \rangle - g^*(y), \quad (3.2a)$$

$$f^*(v, y) = (f_0 + \langle y, F \rangle)^*(v) + g^*(y). \quad (3.2b)$$

Proof of Lemma 3.1 If $x \in \text{dom}(F)$, then

$$\begin{aligned} l(x, y) &= -\sup_{u \in \mathbb{R}^m} \{\langle y, u \rangle - f(x, u)\} = f_0(x) - \sup_{u \in \mathbb{R}^m} \{\langle y, u \rangle - g(F(x) + u)\} \\ &= f_0(x) - \sup_{u \in \mathbb{R}^m} \{\langle y, u - F(x) \rangle - g(u)\} \\ &= f_0(x) + \langle y, F(x) \rangle - g^*(y). \end{aligned}$$

If $x \notin \text{dom}(F)$, then both sides of (3.2a) are $+\infty$.

For (3.2b) we have

$$\begin{aligned}
 f^*(v, y) &= \sup_{x \in \mathbb{R}^n, u \in \mathbb{R}^m} \{ \langle v, x \rangle + \langle y, u \rangle - f(x, u) \} \\
 &= \sup_{x \in \text{dom}(F), u \in \mathbb{R}^m} \{ \langle v, x \rangle + \langle y, u \rangle - f_0(x) - g(F(x) + u) \} \\
 &= \sup_{x \in \text{dom}(F)} \left\{ \langle v, x \rangle - f_0(x) - \inf_{u \in \mathbb{R}^m} \{ g(F(x) + u) - \langle y, u \rangle \} \right\} \\
 &= \sup_{x \in \text{dom}(F)} \left\{ \langle v, x \rangle - f_0(x) - \inf_{u \in \mathbb{R}^m} \{ g(u) - \langle y, u - F(x) \rangle \} \right\} \\
 &= \sup_{x \in \text{dom}(F)} \{ \langle v, x \rangle - f_0(x) - \langle y, F(x) \rangle + g^*(y) \} \\
 &= (f_0 + \langle y, F \rangle)^*(v) + g^*(y)
 \end{aligned}$$

□

5.2. Proofs from Section 4

PROPOSITION 4.1. *Let $g : \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$, $F : \mathbb{R}^n \rightarrow \mathbb{R}^m \cup \{+\infty\}$, $g \circ F : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be given by*

$$g \circ F(x) := \begin{cases} g(F(x)) & x \in \text{dom}(F) \\ +\infty & \text{o.w.} \end{cases}, \quad (4.1)$$

and $f_0 : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be such that $f_0 \equiv 0$ so that $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$ given by (3.1a) becomes

$$f(x, u) := \begin{cases} g(F(x) + u) & x \in \text{dom}(F) \\ +\infty & \text{o.w.} \end{cases} \quad (4.2)$$

and $g \circ F(x) = f(x, 0)$ for all $x \in \mathbb{R}^n$.

If $g \in \Gamma(\mathbb{R}^m)$, then

$$\mathcal{K}_F \cap \mathcal{K}_{g, \text{rng}(F)} \neq \emptyset \Rightarrow g \circ F \text{ is convex} \quad (4.3a)$$

$$\mathcal{K}_F \cap \mathcal{K}_g \neq \emptyset \Rightarrow f \text{ is convex} \quad (4.3b)$$

If $g \in \Gamma_0(\mathbb{R}^m)$ and $\text{dom}(g) \cap \text{rge}(F) \neq \emptyset$, then (4.3b) is an equivalence.

Proof First, (4.3a) follows from [6, Proposition 1.b]. For (4.3b) note that if F is K -convex, then $H : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^m \cup \{+\infty\}$ given by $H(x, u) = F(x) + u$ is also K -convex so $\mathcal{K}_F \subseteq \mathcal{K}_H$. In addition, $\text{rng}(H) = \mathbb{R}^m$ so $\mathcal{K}_g = \mathcal{K}_{g, \text{rng}(H)}$. Hence, $\mathcal{K}_F \cap \mathcal{K}_g \neq \emptyset$ implies $\mathcal{K}_H \cap \mathcal{K}_{g, \text{rng}(H)} \neq \emptyset$, which by (4.3a) implies that $g \circ H = f$ is convex.

Under the additional assumptions on g and F we have that f is proper and for any $\bar{x} \in \mathbb{R}^n$ we have that $f(\bar{x}, \cdot)$ is closed and convex (possibly with $f(\bar{x}, \cdot) \equiv +\infty$). Hence, f is a *dualizing parametrization* as in [15, Definition 11.45]. Then, by [15, Proposition 11.48] and Lemma 3.1 (with $f_0 \equiv 0$) we have that $l : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$ given by $l(x, y) = \langle y, F(x) \rangle - g^*(y)$ is convex in x if and only if $f(x, u)$ is convex in (x, u) . Hence,

$$f(x, u) \text{ is convex in } (x, u) \Leftrightarrow \langle y, F \rangle \text{ is convex } \forall y \in \text{dom}(g^*) \quad (5.1a)$$

$$\Leftrightarrow \langle y, F \rangle \text{ is convex } \forall y \in \text{cl cone dom}(g^*) \quad (5.1b)$$

$$\Leftrightarrow \langle y, F \rangle \text{ is convex } \forall y \in (\text{hzn}(g))^\circ, \quad (5.1c)$$

$$\Leftrightarrow F \text{ is } (-\text{hzn}(g))\text{-convex}, \quad (5.1d)$$

where (5.1b) follows because the set of convex functions is a closed convex cone, (5.1c) follows from [12, Theorem 14.2], and (5.1d) follows from [6, Theorem 3] by noting that $\text{hzn}(g)$ is a closed convex cone (e.g [8, Proposition 55]). Then, by Lemma 4.3 we have that $(-\text{hzn}(g)) \in \mathcal{K}_F \cap \mathcal{K}_g$.

□

LEMMA 4.1. *Let $K \subseteq \mathbb{R}^m$ be a closed convex cone, $g : \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$, $F : \mathbb{R}^n \rightarrow \mathbb{R}^m \cup \{+\infty\}$, $\langle y, F \rangle : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be given by (3.1b) and $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$ given by (4.2).*

If $g \in \Gamma_0(\mathbb{R}^m)$, $K \in \mathcal{K}_F \cap \mathcal{K}_g$, $\text{dom}(g) \cap \text{rge}(F) \neq \emptyset$, and $\langle y, F \rangle \in \Gamma_0(\mathbb{R}^n)$ for all $y \in -K^\circ$, then $f \in \Gamma_0(\mathbb{R}^n \times \mathbb{R}^m)$.

Proof of Lemma 4.1 Follows from the proof of [6, Corollary 1] as follows.

Because $\text{rge}(F) \neq \emptyset$ we have $K\text{-epi}(F) \neq \emptyset$ and [6, Lemma 5] implies $f(x, u) = \psi^*(x, u)$ for $\psi : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$ given by $\psi(v, y) := \sigma_{K\text{-epi}(F)}(v, -y) + g^*(y)$. Under the assumptions, $\sigma_{K\text{-epi}(F)} \in \Gamma_0(\mathbb{R}^n \times \mathbb{R}^m)$ and $g^* \in \Gamma_0(\mathbb{R}^m)$. Then, either $\psi \equiv +\infty$ or $\psi \in \Gamma_0(\mathbb{R}^n \times \mathbb{R}^m)$. However, by [15, Exercise 11.2], if $\psi \equiv +\infty$, then $f \equiv -\infty$, which contradicts $g \in \Gamma_0(\mathbb{R}^m)$ being proper. Hence, $\psi \in \Gamma_0(\mathbb{R}^n \times \mathbb{R}^m)$ and $f \in \Gamma_0(\mathbb{R}^n \times \mathbb{R}^m)$.

□

LEMMA 4.2. *If (4.8) and (4.9a) (or equivalently (4.9b)) holds for g and F , then the general condition $0 \in \text{rint}(U)$ from Theorem 3.1 for $f_0 \equiv 0$ is equivalent to*

$$\exists K \in \mathcal{K}_F \cap \mathcal{K}_g \quad \text{s.t.} \quad \text{rint}(\text{dom}(g)) \cap \text{rint}(F(\text{dom}(F)) + K) \neq \emptyset, \quad (4.10)$$

the weaker condition for piecewise linear-quadratic functions $0 \in U$ from Theorem 3.1 for $f_0 \equiv 0$ is equivalent to

$$\exists K \in \mathcal{K}_F \cap \mathcal{K}_g \quad \text{s.t.} \quad (\text{dom}(g) - K) \cap (F(\text{dom}(F))) \neq \emptyset, \quad (4.11)$$

and condition (4.6) from Theorem 4.1 is equivalent to

$$K \in \mathcal{K}_F \cap \mathcal{K}_g \quad \text{and} \quad \text{rint}(\text{dom}(g) - K) \cap F(\text{rint}(\text{dom}(F))) \neq \emptyset, \quad (4.12)$$

Proof of Lemma 4.2 For $f_0 \equiv 0$ we have that $U = \text{dom}(g) - F(\text{dom}(F))$. In addition, by [6, Lemma 3] if $K \in \mathcal{K}_g$, then $\text{dom}(g) = \text{dom}(g) - K$. Hence,

$$U = (\text{dom}(g) - K) - F(\text{dom}(F)) = \text{dom}(g) - (F(\text{dom}(F)) + K),$$

so $0 \in U$ is equivalent to (4.11). We also have that if $K \in \mathcal{K}_F$, then $\text{rng}(F) + K$ is convex because it is the projection of the K -epigraph of F onto the z variables. Then, by convexity of $\text{dom}(g)$ and [12, Corollary 6.6.2] we have that

$$\text{rint}(U) = \text{rint}(\text{dom}(g)) - \text{rint}(F(\text{dom}(F)) + K)$$

and hence $0 \in \text{rint}(U)$ is equivalent to (4.10).

Condition (4.12) is simply a re-ordering of (4.6). \square

References

- [1] Dimitri Bertsekas. *Convex optimization theory*, volume 1. Athena Scientific, 2009.
- [2] Radu Ioan Boț. *Conjugate duality in convex optimization*, volume 637. Springer Science & Business Media, 2009.
- [3] Radu Ioan Boț, Sorin-Mihai Grad, and Gert Wanka. New constraint qualification and conjugate duality for composed convex optimization problems. *Journal of Optimization Theory and Applications*, 135(2):241–255, 2007.
- [4] Radu Ioan Boț, Sorin-Mihai Grad, and Gert Wanka. A new constraint qualification for the formula of the subdifferential of composed convex functions in infinite dimensional spaces. *Mathematische Nachrichten*, 281(8):1088–1107, 2008.

- [5] Radu Ioan Bot, Sorin-Mihai Grad, and Gert Wanka. Generalized moreau–rockafellar results for composed convex functions. *Optimization*, 58(7):917–933, 2009.
- [6] James V Burke, Hoheisel Tim, and Quang V Nguyen. A study of convex convex-composite functions via infimal convolution with applications. *Mathematics of Operations Research*, 46(4):1324–1348, 2021.
- [7] Christophe Combari, Mohamed Laghdif, and Lionel Thibault. Sous-différentiels de fonctions convexes composées. *Ann. Sci. Math. Québec*, 18(2):119–148, 1994.
- [8] Armand Gissler and Tim Hoheisel. A note on the K -epigraph. *Optimization*, 72(9):2251–2285, 2023.
- [9] Jean-Baptiste Hiriart-Urruty. A note on the legendre-fenchel transform of convex composite functions. In *Nonsmooth Mechanics and Analysis: Theoretical and Numerical Advances*, pages 35–46. Springer, 2006.
- [10] Jean-Baptiste Hiriart-Urruty and Claude Lemaréchal. *Fundamentals of convex analysis*. Springer Science & Business Media, 2004.
- [11] Teemu Pennanen. Graph-convex mappings and K -convex functions. *Journal of Convex Analysis*, 6(2):235–266, 1999.
- [12] R. Tyrrell Rockafellar. *Convex Analysis*. Princeton Landmarks in Mathematics and Physics. Princeton University Press, 2015.
- [13] R Tyrrell Rockafellar. Augmented lagrangians and hidden convexity in sufficient conditions for local optimality. *Mathematical Programming*, 198(1):159–194, 2023.
- [14] R Tyrrell Rockafellar. Convergence of augmented lagrangian methods in extensions beyond nonlinear programming. *Mathematical Programming*, 199(1):375–420, 2023.
- [15] R Tyrrell Rockafellar and Roger J-B Wets. *Variational analysis*, volume 317. Springer Science & Business Media, 2009.

Appendix A: Proofs and Examples from Section 2

A.1. Proof of Theorem 2.1, Proposition 2.1 and Theorem 2.2 To prove Theorem 2.1, Proposition 2.1 and Theorem 2.2 we rely on the following lemmas.

LEMMA A.1. *Let $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$. Then h^* is convex and*

$$h \geq \text{co}(h) \geq \text{cl co}(h) \geq h^{**}. \quad (\text{A.1})$$

Equality holds throughout (A.1) if $h \in \Gamma_0(\mathbb{R}^m)$ or $h \equiv +\infty$. If $h \in \Gamma(\mathbb{R}^m)$, then both $h^ \in \Gamma_0(\mathbb{R}^m)$, $h^{**} \in \Gamma_0(\mathbb{R}^m)$ and $\text{cl}(h) = h^{**}$.*

Proof Direct from [15, Theorem 11.1 and Exercise 11.2]. \square

LEMMA A.2. For $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$, $\bar{v} \in \mathbb{R}^n$, and $\bar{u} \in \mathbb{R}^m$ let $p_{\bar{v}} : \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$ and $q_{\bar{u}} : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be given by

$$p_{\bar{v}}(\bar{u}) := \inf_{x \in \mathbb{R}^n} \{f(x, \bar{u}) - \langle \bar{v}, x \rangle\}$$

$$q_{\bar{u}}(\bar{v}) := \inf_{y \in \mathbb{R}^m} \{f^*(\bar{v}, y) - \langle y, \bar{u} \rangle\}.$$

Then, $q_{\bar{u}}$ is convex and for all $\bar{x} \in \mathbb{R}^n$, $\bar{y} \in \mathbb{R}^m$, $\bar{v} \in \mathbb{R}^n$, and $\bar{u} \in \mathbb{R}^m$ we have

$$p_{\bar{v}}^*(\bar{y}) = f^*(\bar{v}, \bar{y}) \quad (\text{A.2a})$$

$$p_{\bar{v}}^{**}(\bar{u}) = \sup_{y \in \mathbb{R}^m} \{\langle \bar{u}, y \rangle - f^*(\bar{v}, y)\} = -q_{\bar{u}}(\bar{v}) \quad (\text{A.2b})$$

$$q_{\bar{u}}^*(\bar{x}) = f^{**}(\bar{x}, \bar{u}) \leq f(\bar{x}, \bar{u}) \quad (\text{A.2c})$$

$$q_{\bar{u}}^{**}(\bar{v}) \geq \sup_{x \in \mathbb{R}^n} \{\langle \bar{v}, x \rangle - f(x, \bar{u})\} = -p_{\bar{v}}(\bar{u}). \quad (\text{A.2d})$$

If $f \in \Gamma_0(\mathbb{R}^n \times \mathbb{R}^m)$, then $p_{\bar{v}}$ is also convex and equality holds in (A.2c) and (A.2d).

Proof Convexity of $q_{\bar{u}}$ follows from [15, Proposition 2.22] by noting that by Lemma A.1 f^* is convex.

The equations in (A.2a) and (A.2c) follow by expanding the definitions of $p_{\bar{v}}^*$ and $q_{\bar{u}}^*$ and the inequality in (A.2c) follows from Lemma A.1. We then directly get (A.2b) and (A.2d).

Under the additional conditions on f convexity of $p_{\bar{v}}$ follows from [15, Proposition 2.22] and the additional equalities follow from Lemma A.1 applied to f . \square

LEMMA A.3. Let $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be a convex function, $\bar{x} \in \mathbb{R}^n$ and $\bar{v} \in \mathbb{R}^n$. Then,

$$\arg \max_{x \in \mathbb{R}^n} \{\langle \bar{v}, x \rangle - h(x)\} = \partial h^*(\bar{v})$$

$$\arg \max_{v \in \mathbb{R}^n} \{\langle v, \bar{x} \rangle - h^*(v)\} = \partial h(\bar{x})$$

Proof Follows from [15, Theorem 11.3] and [15, Exercise 11.2]. \square

THEOREM 2.1. *For any $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$ and $(\bar{v}, \bar{u}) \in \mathbb{R}^n \times \mathbb{R}^m$ we have that $q_{\bar{u}}$ is convex,*

$$f^*(\bar{v}, \bar{y}) - \langle \bar{u}, \bar{y} \rangle = \sup_{x \in \mathbb{R}^n} \{ \langle \bar{v}, x \rangle - \langle \bar{u}, \bar{y} \rangle - l(x, \bar{y}) \} \quad \forall \bar{y} \in \mathbb{R}^m$$

and $q_{\bar{u}}(\bar{v}) := \inf_{y \in \mathbb{R}^m} \sup_{x \in \mathbb{R}^n} \{ \langle \bar{v}, x \rangle - \langle \bar{u}, y \rangle - l(x, y) \}$.

In addition,

$$\text{opt}_P(\bar{u}, \bar{v}) = p_{\bar{v}}(\bar{u}) \geq \text{cl co } p_{\bar{v}}(\bar{u}) \geq -q_{\bar{u}}(\bar{v}) = \text{opt}_D(\bar{u}, \bar{v}), \quad (2.5)$$

$$-\text{opt}_P(\bar{u}, \bar{v}) = -p_{\bar{v}}(\bar{u}) \leq \text{cl } q_{\bar{u}}(\bar{v}) \leq q_{\bar{u}}(\bar{v}) = -\text{opt}_D(\bar{u}, \bar{v}). \quad (2.6)$$

Finally, for any $(\bar{v}, \bar{u}) \in \mathbb{R}^n \times \mathbb{R}^m$ and $(\bar{x}, \bar{y}) \in \mathbb{R}^n \times \mathbb{R}^m$

$$\bar{x} \in P(\bar{u}, \bar{v}), \quad \bar{y} \in Q(\bar{u}, \bar{v}), \quad \text{and} \quad p_{\bar{v}}(\bar{u}) = -q_{\bar{u}}(\bar{v}) \quad (2.7)$$

is equivalent to

$$f(\bar{x}, \bar{u}) - \langle \bar{v}, \bar{x} \rangle = l(\bar{x}, \bar{y}) + \langle \bar{u}, \bar{y} \rangle - \langle \bar{v}, \bar{x} \rangle = \langle \bar{u}, \bar{y} \rangle - f^*(\bar{v}, \bar{y}) \in \mathbb{R}. \quad (2.8)$$

Proof of Theorem 2.1

We begin by noting that for all $\bar{v} \in \mathbb{R}^n$ and $\bar{y} \in \mathbb{R}^m$ we have

$$f^*(\bar{v}, \bar{y}) = \sup_{x \in \mathbb{R}^n} \{ \langle \bar{v}, x \rangle - l(x, \bar{y}) \} \quad (\text{A.3})$$

which yields the last equation in (2.5) and (2.6).

Then (2.5) and (2.6) follow from Lemma A.2 and Lemma A.1 (applied to the functions $p_{\bar{v}}$ and $q_{\bar{u}}$ defined in Lemma A.2).

The equivalence between (2.7), (2.8) follows by noting that for any $(\bar{x}, \bar{y}) \in \mathbb{R}^n \times \mathbb{R}^m$

$$f(\bar{x}, \bar{u}) - \langle \bar{v}, \bar{x} \rangle \geq (f(\bar{x}, \cdot))^{\ast\ast}(\bar{u}) - \langle \bar{v}, \bar{x} \rangle \quad (\text{A.4a})$$

$$= \sup_{y \in \mathbb{R}^m} \{ l(\bar{x}, y) + \langle \bar{u}, y \rangle - \langle \bar{v}, \bar{x} \rangle \} \quad (\text{A.4b})$$

$$\geq l(\bar{x}, \bar{y}) + \langle \bar{u}, \bar{y} \rangle - \langle \bar{v}, \bar{x} \rangle \quad (\text{A.4c})$$

$$\geq \inf_{x \in \mathbb{R}^n} \{ l(x, \bar{y}) + \langle \bar{u}, \bar{y} \rangle - \langle \bar{v}, x \rangle \} \quad (\text{A.4d})$$

$$= \langle \bar{u}, \bar{y} \rangle - f^*(\bar{v}, \bar{y}) \quad (\text{A.4e})$$

where the first inequality follows from Lemma A.1, the first equality follows from $l(x, y) = -(f(x, \cdot))^*(y)$ and the last equality follows from (A.3). \square

PROPOSITION 2.1. *For any proper $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$ such that $f(x, u)$ is closed and convex in u , and for any $(\bar{v}, \bar{u}) \in \mathbb{R}^n \times \mathbb{R}^m$ we have*

$$f(\bar{x}, \bar{u}) - \langle \bar{v}, \bar{x} \rangle = \sup_{y \in \mathbb{R}^m} \{l(\bar{x}, y) + \langle \bar{u}, y \rangle - \langle \bar{v}, \bar{x} \rangle\} \quad \forall \bar{x} \in \mathbb{R}^n$$

and $p_{\bar{v}}(\bar{u}) := \inf_{x \in \mathbb{R}^n} \sup_{y \in \mathbb{R}^m} \{l(x, y) + \langle \bar{u}, y \rangle - \langle \bar{v}, x \rangle\}$.

In addition, (2.8) is equivalent to

$$\bar{x} \in \arg \min_{x \in \mathbb{R}^n} \{l(x, \bar{y}) + \langle \bar{u}, \bar{y} \rangle - \langle \bar{v}, x \rangle\}, \quad (2.9a)$$

$$\bar{y} \in \arg \max_{y \in \mathbb{R}^m} \{l(\bar{x}, y) + \langle \bar{u}, y \rangle - \langle \bar{v}, \bar{x} \rangle\}. \quad (2.9b)$$

Proof of Proposition 2.1

All results follow from (A.4) by noting that if f is proper and $f(x, u)$ is closed and convex in u , then $f(x, \cdot) \equiv +\infty$ or $f(x, u)$ is proper, closed and convex in u so by Lemma A.1 we have $f(\bar{x}, \bar{u}) = (f(\bar{x}, \cdot))^{\ast\ast}(\bar{u})$. \square

THEOREM 2.2. *If $f \in \Gamma(\mathbb{R}^n \times \mathbb{R}^m)$ and $f(x, u)$ is closed in u , then $l(x, y)$ is convex in x and concave in y and (2.9) is equivalent to*

$$0 \in \partial(l(\cdot, \bar{y}) + \langle \bar{u}, \bar{y} \rangle - \langle \bar{v}, \cdot \rangle)(\bar{x}) \quad (2.10a)$$

$$0 \in \partial(-l(\bar{x}, \cdot) - \langle \bar{u}, \cdot \rangle + \langle \bar{v}, \bar{x} \rangle)(\bar{y}). \quad (2.10b)$$

In addition, if $f \in \Gamma_0(\mathbb{R}^n \times \mathbb{R}^m)$ we have

1. **Optimal sets:** For all $(\bar{v}, \bar{u}) \in \mathbb{R}^n \times \mathbb{R}^m$ we have that $p_{\bar{v}}$ and $q_{\bar{u}}$ are convex, and

$$P(\bar{u}, \bar{v}) = \partial q_{\bar{u}}(\bar{v}) \quad (2.11a)$$

$$Q(\bar{u}, \bar{v}) = \partial p_{\bar{v}}(\bar{u}). \quad (2.11b)$$

2. **Primal-asymptotic strong duality:** If $p_{\bar{v}}$ is proper, then $p_{\bar{v}}(\bar{u}) \geq \text{cl } p_{\bar{v}}(\bar{u}) = -q_{\bar{u}}(\bar{v})$ for all $\bar{u} \in \mathbb{R}^m$.

3. **Primal-uniform strong duality:** If $p_{\bar{v}}$ is proper and closed, then $p_{\bar{v}}(\bar{u}) = -q_{\bar{u}}(\bar{v})$ for all $\bar{u} \in \mathbb{R}^m$.

4. **Dual-asymptotic strong duality:** If $q_{\bar{u}}$ is proper, then $-p_{\bar{v}}(\bar{u}) = \text{cl } q_{\bar{u}}(\bar{v})$ for all $\bar{v} \in \mathbb{R}^n$.

5. **Dual-uniform strong duality:** If $q_{\bar{u}}$ is proper and closed, then $-p_{\bar{v}}(\bar{u}) = q_{\bar{u}}(\bar{v})$ for all $\bar{v} \in \mathbb{R}^n$.

Proof of Theorem 2.2

Convexity of $q_{\bar{u}}$ follows from Lemma A.2.

The convexity statements for l follow from [15, Proposition 11.48] and the equivalence between (2.9) and (2.10) follows from Lemma A.3.

The statements on $p_{\bar{v}}$ and $q_{\bar{u}}$ follow from Lemmas A.2 and A.3.

The remainder of the results follow from Lemma A.2, Lemma A.1 (applied to the functions $p_{\bar{v}}$ and $q_{\bar{u}}$). \square

A.2. Qualification conditions Most sufficient conditions for the qualification conditions in Theorem 2.2 are based on the domains of functions $p_{\bar{v}}$ and $q_{\bar{u}}$. The following straightforward lemma notes that for such conditions we can restrict our attention to the cases $\bar{v} = 0$ and $\bar{u} = 0$.

LEMMA A.4. *Let $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$, $\bar{v} \in \mathbb{R}^n$, $\bar{u} \in \mathbb{R}^m$, and $p_{\bar{v}} : \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$, $q_{\bar{u}} : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be the functions given by (2.4). Then,*

$$\text{dom}(p_0) = \text{dom}(p_{\bar{v}}) \quad \forall \bar{v} \in \mathbb{R}^n$$

$$\text{dom}(q_0) = \text{dom}(q_{\bar{u}}) \quad \forall \bar{u} \in \mathbb{R}^m.$$

A.2.1. Conditions for asymptotic strong duality For $f \in \Gamma(\mathbb{R}^n \times \mathbb{R}^m)$, both $p_{\bar{v}}$ and $q_{\bar{u}}$ are guaranteed to not be identically $+\infty$. Exploiting some dual relationships between these functions (e.g. Lemma A.2), we can ensure one of the function avoids taking the value $-\infty$ through conditions on the domain of the other one as follows.

LEMMA A.5. *Let $f \in \Gamma(\mathbb{R}^n \times \mathbb{R}^m)$, $\bar{v} \in \mathbb{R}^n$, $\bar{u} \in \mathbb{R}^m$, and $p_{\bar{v}} : \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$, $q_{\bar{u}} : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be the functions given by (2.4).*

- If $\bar{v} \in \text{dom}(q_0)$ (or equivalently if $f^*(\bar{v}, \cdot) \not\equiv +\infty$), then $p_{\bar{v}}$ is proper, and
- if $\bar{u} \in \text{dom}(p_0)$ (or equivalently if $f(\cdot, \bar{u}) \not\equiv +\infty$), then $q_{\bar{u}}$ is proper

Proof We start by noting that $p_{\bar{v}} \not\equiv +\infty$ for all $\bar{v} \in \mathbb{R}^n$ and $q_{\bar{u}} \not\equiv +\infty$ for all $\bar{u} \in \mathbb{R}^m$. This follows from f and f^* being proper under the assumptions on f and Lemma A.1.

For the remaining statements, first note that if $f^*(\bar{v}, \cdot) \not\equiv +\infty$, then $q_{\bar{u}}(\bar{v}) < +\infty$ for all $\bar{u} \in \mathbb{R}^m$. Similarly, if $f(\cdot, \bar{u}) \not\equiv +\infty$, then $p_{\bar{v}}(\bar{u}) < +\infty$ for all $\bar{v} \in \mathbb{R}^n$. In addition, by Lemmas A.1 and A.2, we have that $p_{\bar{v}}(\bar{u}) \geq -q_{\bar{u}}(\bar{v})$ and $q_{\bar{u}}(\bar{v}) \geq -p_{\bar{v}}(\bar{u})$ for all $\bar{v} \in \mathbb{R}^n$ and $\bar{u} \in \mathbb{R}^m$. Hence, if $f^*(\bar{v}, \cdot) \not\equiv +\infty$, then $p_{\bar{v}}(\bar{u}) > -\infty$ for all $\bar{u} \in \mathbb{R}^m$ and if $f(\cdot, \bar{u}) \not\equiv +\infty$, then $q_{\bar{u}}(\bar{v}) > -\infty$ for all $\bar{v} \in \mathbb{R}^n$. The result follows because we already have that $p_{\bar{v}}$ and $q_{\bar{u}}$ are never identically $+\infty$. \square

A.2.2. Conditions for non-asymptotic strong duality A simple Slater-like condition for non-asymptotic strong duality can be obtained by replacing dom by the interior of dom in Lemma A.5. However, Proposition A.1 below shows we can obtain slightly weaker conditions in general and much weaker conditions for PWLQ functions.

PROPOSITION A.1. *Let $f \in \Gamma_0(\mathbb{R}^n \times \mathbb{R}^m)$, $\bar{v} \in \mathbb{R}^n$, $\bar{u} \in \mathbb{R}^m$, and $p_{\bar{v}} : \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$, $q_{\bar{u}} : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be the functions given by (2.4).*

Then $q_{\bar{u}} \in \Gamma_0(\mathbb{R}^n)$ if any of the following holds

- $\bar{u} \in \text{rint dom}(p_0)$, or
- $\bar{u} \in \text{dom}(p_0)$ and f is PWLQ.

Proof First note that

$$L(C) \subseteq \text{epi}(q_{\bar{u}}) \subseteq \text{cl}(L(C)) \quad (\text{A.5})$$

for $C := \{(v, y, z) \in \mathbb{R}^{n+m+1} : f^*(v, y) - \langle \bar{u}, y \rangle \leq z\}$ and the linear transformation $L : \mathbb{R}^{n+m+1} \rightarrow \mathbb{R}^{n+1}$ given by $L(v, y, z) = (v, z)$ whose nullspace is $N(L) := \{(v, y, z) \in \mathbb{R}^{n+m+1} : v = 0, z = 0\}$ (e.g. [1, Proposition 3.3.1]). In addition, by Lemma A.1 we have that $f^* \in \Gamma_0(\mathbb{R}^n \times \mathbb{R}^m)$, so C is closed and convex. By [15, Exercise 3.29] we have that

$$C_\infty \cap N(L) = \{(v, y, z) \in \mathbb{R}^{n+m+1} : f^{*\infty}(0, y) - \langle \bar{u}, y \rangle \leq 0, v = 0, z = 0\}.$$

We claim that

$$C_\infty \cap N(L) \subseteq (-(C_\infty \cap N(L))) \cap (C_\infty \cap N(L)) \quad (\text{A.6})$$

is equivalent to $\bar{u} \in \text{rint dom}(p_0)$. To show this first note that $\bar{u} \in \text{rint dom}(p_0)$ is equivalent to having $y \neq 0$ imply

$$\sigma_{\text{dom}(p_0)}(y) > \langle \bar{u}, y \rangle, \quad \text{or} \quad -\sigma_{\text{dom}(p_0)}(-y) = \sigma_{\text{dom}(p_0)}(y) = \langle \bar{u}, y \rangle.$$

In addition, by [15, Theorem 11.5] $\sigma_{\text{dom}(p_0)}(y) = \sigma_{\text{dom}(f)}(0, y) = f^{*\infty}(0, y)$. Hence, $\bar{u} \in \text{rint dom}(p_0)$ is also equivalent to having $y \neq 0$ imply

$$f^{*\infty}(0, y) - \langle \bar{u}, y \rangle > 0, \quad \text{or} \quad f^{*\infty}(0, y) - \langle \bar{u}, y \rangle = f^\infty(0, -y) - \langle \bar{u}, -y \rangle = 0,$$

which in turn is equivalent to (A.6)

Now by [1, Proposition 1.4.13] we have that $L(C)$ is closed if $C_\infty \cap N(L) \subseteq (-(C_\infty \cap N(L))) \cap (C_\infty \cap N(L)) \subseteq (-C_\infty) \cap C_\infty$. Hence, the first sufficient condition implies $L(C)$ is closed. By (A.5), this condition also implies $q_{\bar{u}}$ is closed. In addition, from the equivalence between (A.6) and $\bar{u} \in \text{rint dom}(p_0)$ and Lemma A.5 we have that $q_{\bar{u}}$ is proper. Convexity of $q_{\bar{u}}$ follows from convexity of f^* by [15, Proposition 2.22].

The second sufficient condition follows from [15, Proposition 10.21, Theorem 11.14 and Proposition 11.32]. \square

PROPOSITION A.2. *Let $f \in \Gamma_0(\mathbb{R}^n \times \mathbb{R}^m)$, $\bar{v} \in \mathbb{R}^n$, $\bar{u} \in \mathbb{R}^m$, and $p_{\bar{v}} : \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$, $q_{\bar{u}} : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be the functions given by (2.4).*

Then $p_{\bar{v}} \in \Gamma_0(\mathbb{R}^m)$ if any of the following holds

- $\bar{v} \in \text{rint dom}(q_0)$, or
- $\bar{v} \in \text{dom}(q_0)$ and f is PWLQ.

Proof Analogous to the proof of Proposition A.1. \square

A.2.3. Conditions for optimal value attainment Non-emptiness of the sub-differentials in (2.11) can be ensured through standard results such as the following lemma.

LEMMA A.6. *Let $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be a convex function and $\bar{x} \in \mathbb{R}^n$. Then, $\partial h(\bar{x}) \neq \emptyset$ if any of the following holds*

- $\bar{x} \in \text{rint dom}(h)$ and $h(\bar{x}) > -\infty$, or
- $\bar{x} \in \text{dom}(h)$ and h is piecewise linear-quadratic.

Proof For the first case, we begin by noting that if $\bar{x} \in \text{rint dom}(h)$ and $h(\bar{x}) > -\infty$, then h is proper. Indeed, $h(\bar{x}) \in \mathbb{R}$ so $h \not\equiv +\infty$. In addition, if there exist $\bar{x}' \in \text{dom}(h)$ such that $h(\bar{x}') = -\infty$, then by convexity of h and $\bar{x} \in \text{rint dom}(h)$, there exist $\lambda > 0$ such that

$$h(\bar{x}) \leq \lambda h\left(\frac{\bar{x} + (\lambda - 1)\bar{x}'}{\lambda}\right) + (1 - \lambda)h(\bar{x}') = -\infty,$$

which contradicts $h(\bar{x}) > -\infty$. The result then follows from [12, Theorem 23.4].

The second case follows from [15, Proposition 10.21]. \square

A.2.4. Slater-like conditions

COROLLARY 2.1. [Primal Slater Condition] Let $f \in \Gamma_0(\mathbb{R}^n \times \mathbb{R}^m)$, $\bar{v} \in \mathbb{R}^n$, $\bar{u} \in \mathbb{R}^m$, and $p_{\bar{v}} : \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$, $q_{\bar{u}} : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be the functions given by (2.4). Then,

- $q_{\bar{u}}$ is proper if $\bar{u} \in \text{dom}(p_0)$,
- $q_{\bar{u}}$ is closed if $\bar{u} \in \text{rint dom}(p_0)$, or $\bar{u} \in \text{dom}(p_0)$ and f is PWLQ, and
- $\partial p_{\bar{v}}(\bar{u}) \neq \emptyset$ if $\bar{u} \in \text{rint dom}(p_0)$ and $p_{\bar{v}}(\bar{u}) > -\infty$, or $\bar{u} \in \text{dom}(p_0)$ and f is PWLQ,

In particular, if either $\bar{u} \in \text{rint dom}(p_0)$, or $\bar{u} \in \text{dom}(p_0)$ and f is PWLQ, then for any $\bar{v} \in \mathbb{R}^n$ we have

$$\exists \bar{y} \in \mathbb{R}^m \quad s.t \quad \langle \bar{u}, \bar{y} \rangle - f^*(\bar{v}, \bar{y}) = -q_{\bar{u}}(\bar{v}) = p_{\bar{v}}(\bar{u}) \in \mathbb{R} \cup \{-\infty\}. \quad (2.12)$$

Proof of Corollary 2.1

By [15, Theorem 11.14 and Proposition 11.32] we have that if f is PWLQ then so are $p_{\bar{v}}$ and $q_{\bar{u}}$. The bullets follow from Proposition A.1, Lemmas A.5 and A.6. Theorem 2.2 (and $q_{\bar{u}}$ being proper) then implies $-q_{\bar{u}}(\bar{v}) = p_{\bar{v}}(\bar{u}) \in \mathbb{R} \cup \{-\infty\}$ with $-q_{\bar{u}}(\bar{v}) \in \mathbb{R} \cup \{-\infty\}$. If $p_{\bar{v}}(\bar{u}) > -\infty$, then (2.12) holds for any $\bar{y} \in \partial p_{\bar{v}}(\bar{u})$ by Theorem 2.2. If $q_{\bar{u}}(\bar{v}) = -p_{\bar{v}}(\bar{u}) = +\infty$, then (2.12) holds for any $\bar{y} \in \mathbb{R}^m$ by the definition of $q_{\bar{u}}$ in (2.4b). \square

COROLLARY 2.2. [Dual Slater Condition] Let $f \in \Gamma_0(\mathbb{R}^n \times \mathbb{R}^m)$, $\bar{v} \in \mathbb{R}^n$, $\bar{u} \in \mathbb{R}^m$, and $p_{\bar{v}} : \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$, $q_{\bar{u}} : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be the functions given by (2.4). Then,

- $p_{\bar{v}}$ is proper if $\bar{v} \in \text{dom}(q_0)$,
- $p_{\bar{v}}$ is closed if $\bar{v} \in \text{rint dom}(q_0)$, or $\bar{v} \in \text{dom}(q_0)$ and f is PWLQ, and
- $\partial q_{\bar{u}}(\bar{v}) \neq \emptyset$ if $\bar{v} \in \text{rint dom}(q_0)$ and $q_{\bar{u}}(\bar{v}) > -\infty$, or $\bar{v} \in \text{dom}(q_0)$ and f is PWLQ.

In particular, if either $\bar{v} \in \text{rint dom}(q_0)$, or $\bar{v} \in \text{dom}(q_0)$ and f is PWLQ, then for any $\bar{u} \in \mathbb{R}^m$ we have

$$\exists \bar{x} \in \mathbb{R}^n \quad s.t \quad f(\bar{x}, \bar{u}) - \langle \bar{v}, \bar{x} \rangle = p_{\bar{v}}(\bar{u}) = -q_{\bar{u}}(\bar{v}) \in \mathbb{R} \cup \{+\infty\}. \quad (2.13)$$

Proof of Corollary 2.2

Analogous to the proof of Corollary 2.1 except for using Proposition A.2 instead of Proposition A.1. \square